THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Shared mutations in a novel glutaredoxin repressor of multicellular trichome fate underlie parallel evolution of Antirrhinum species

Citation for published version:
Tan, Y, Barnbrook, M, Wilson, Y, Molnar, A, Bukys, A & Hudson, A 2020, 'Shared mutations in a novel glutaredoxin repressor of multicellular trichome fate underlie parallel evolution of Antirrhinum species', *Current Biology*, vol. 30, no. 8, pp. 1357-1366.E4. https://doi.org/10.1016/j.cub.2020.01.060

Digital Object Identifier (DOI):
10.1016/j.cub.2020.01.060

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Current Biology

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 05. May. 2023

Title: Shared mutations in a novel glutaredoxin repressor of multicellular trichome fate underlie parallel evolution of *Antirrhinum* species

Authors: Ying Tan, Matthew Barnbrook, Yvette Wilson, Attila Molnár, Alfredas Bukys, Andrew Hudson*.

Affiliation: University of Edinburgh, Institute of Molecular Plant Sciences, Max Born Crescent, Edinburgh, EH9 3BF, UK.

* Lead Contact/corresponding author, email andrew.hudson@ed.ac.uk

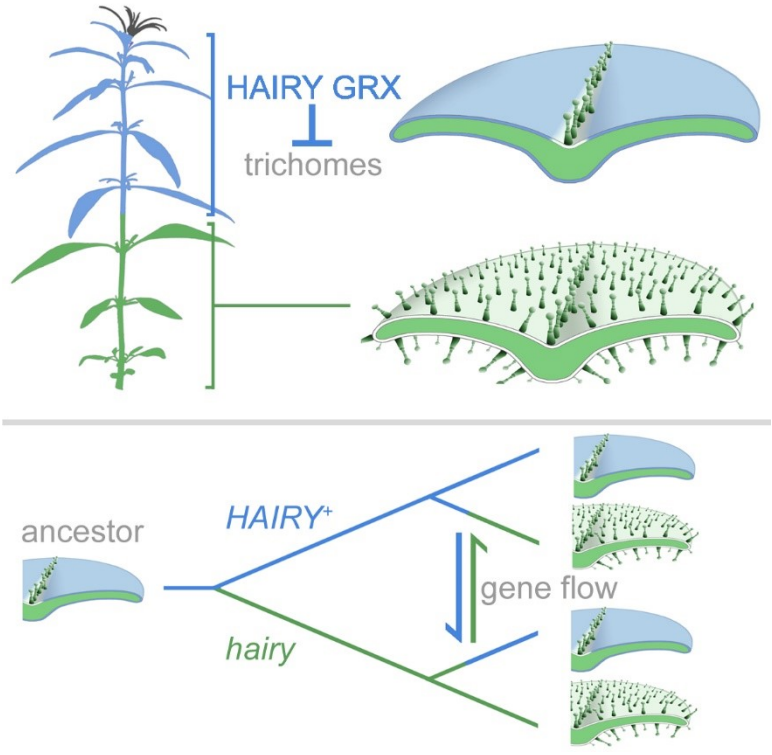
Highlights

- *Hairy* encodes a trichome-repressing glutaredoxin.
- Phased *Hairy* expression determines which tissues are bald.
- Hairy alpine *Antirrhinum* species lack Hairy function.
- Old *hairy* mutations were used in evolution of new alpine species.

eTOC blurb

Tan *et al*. identify a glutaredoxin, encoded by the *Antirrhinum* Hairy gene, that represses multicellular hair (trichome) formation in response to developmental phase. Loss of Hairy was involved in the early origin of alpine species with dense trichomes and old alleles were recycled in parallel evolution of alpine or lowland species.

GRAPHICAL ABSTRACT



SUMMARY

Most angiosperms produce trichomes--epidermal hairs that have protective or more specialised roles. Trichomes are multicellular in almost all species and, in the majority, secretory. Despite the importance of multicellular trichomes for plant protection and as a source of high-value products, the mechanisms that control their development are only poorly understood. Here we investigate the control of multicellular trichome patterns using natural variation within the genus *Antirrhinum* (snapdragons), which has evolved hairy alpine-adapted species or lowland species with a restricted trichome pattern multiple times in parallel. We find that a single gene, *Hairy* (*H*), which is needed to repress trichome fate, underlies variation in trichome patterns between all *Antirrhinum* species except one. We show that *H* encodes a novel epidermis-specific glutaredoxin and that the pattern of trichome distribution within individuals reflects the location of *H* expression. Phylogenetic and functional tests suggest that *H* gained its trichome-repressing role late in the history of eudicots and that the ancestral *Antirrhinum* had an active *H* gene and restricted trichome distribution. Loss of *H* function was involved in an early divergence of alpine and lowland *Antirrhinum* lineages and the alleles underlying this split were later reused in parallel evolution of alpines from lowland ancestors, and *vice versa*. We also find evidence for an evolutionary reversal from a widespread to restricted trichome distribution involving a suppressor mutation and for a pleiotropic effect of *H* on plant growth that might constrain the evolution of trichome pattern.

INTRODUCTION

Angiosperms produce epidermal hairs (trichomes) that are mostly associated with adaptation to herbivores or abiotic factors such as UV radiation [1, 2]. In almost all species they are multicellular, and in many have a similar structure of apical secretory cells supported by a stalk [3], which suggests homology and a single, ancient origin. Because trichomes protect plants and their secretions are the source of economically important compounds, including pharmaceuticals and flavours, regulation of their development is a target for crop improvement and biotechnology [4-6].

Although associated with beneficial roles in protection, trichomes often differ in distribution within individual plants (e.g [7, 8] or between close relatives (e.g.,[9]). This suggests either that there are fitness costs of trichomes that trade off against protection or that changes in trichome distribution are developmentally constrained. These possibilities are difficult to distinguish for multicellular trichomes because control of their development is relatively poorly understood. Their formation in diverse eudicots is known to require an HD-Zip IV transcription factor, supporting their homology [10-15]. They also require a zinc-finger protein in tomato [16] and Mixta-like MYB transcription factors in cotton and *Artemisia annua* [17, 18], though *Mixta-like* genes can induce ectopic trichomes in other species when mis-expressed in transgenic plants (e.g., [19]). It also seems likely that the mechanism specifying

multicellular trichome fate is not shared with the unicellular trichomes of the Brassicaceae, including Arabidopsis, in which patterning of the root epidermis and trichome fate specification involve common genes (reviewed by [20]). This is consistent with parallel evolution of the two trichome types. In this scenario, the Brassicacea lost the ancestral multicellular trichomes and replaced them with unicellular trichomes specified by a different regulatory network, possibly co-opted from metabolic control [21]. Parallel evolution is further supported by requirement for HD-ZIP IV genes that were recruited independently from paralogues that diverged before the origin of seed plants [22, 23]. Ancestral epidermal expression can explain why members of this gene family were predisposed to being recruited in parallel evolution [24]. However, unicellular trichomes might have evolved multiple times in different lineages. In cotton, for example, homology of fibres (unicellular seed trichomes) and multicellular leaf trichomes is suggested by their common requirement for *Mixta-like* function [17], suggesting that cotton fibres evolved independently from unicellular trichomes of Brassicaceae.

One way to reveal more of the mechanism controlling multicellular trichome fate and of the constraints on its evolution is to exploit natural variation. Here we apply this approach using the genus *Antirrhinum* (snapdragons), which comprises ~25 species differing in multicellular trichome morphology and the pattern of trichome distribution within individuals. We show that a single gene, *Hairy* (*H*), accounts for the restricted distribution of trichomes in all lowland *Antirrhinum* species except one. *H*, which encodes a glutaredoxin, is needed to suppress trichomes fate and is expressed in epidermal cells in response to developmental phase. Combined phylogenetic and functional analysis suggests that *H* was recruited to suppress trichome formation late in the evolutionary history of eudicots, that its loss was involved in the early divergence of a hairy alpine-adapted *Antirrhinum* lineage from a lowland lineage with restricted trichomes and that the *H* alleles underlying this divergence were later reused in parallel evolution of further alpine and lowland species.

RESULTS

Antirrhinum species have traditionally been divided into three morphological subsections that correlate with ecology--small, prostrate alpine species in subsection *Kickxiella* and large, upright lowland species in subsections *Antirrhinum* and *Streptosepalum* [25-27]. However, molecular phylogenies support parallel evolution of *Kickxiella* morphology; first in an early divergence of a basal *Kickxiella* lineage from lowland subsection *Antirrhinum* and later from within subsection *Antirrhinum* in Southeast Spain (Figure 1A). Similarly, the large lowland *Streptosepalum* species evolved within the basal *Kickxiella* alpine lineage. It is not currently clear whether this parallel evolution involved introgression of older alleles, sorting of ancestral variation, or mutation *de novo* [27].

The many characters that distinguish morphological subsections include the pattern of trichome distribution. Under the same environmental conditions, all species produce glandular multicellular trichomes from leaves and internodes below metamer 4 (m4, where cotyledons and the internode

above them are m1; Figure 1D, F; Figure S1G-K). In common with many alpine plants, *Kickxiella* species remain densely hairy throughout development, although they differ from each other in trichome morphology from m4 (e.g., Figure 1J, K; Figure S1). In contrast, leaf blades and stems above m4 lack trichomes in lowland subsections *Antirrhinum* and *Streptosepalum*, though glandular trichomes are retained on adaxial leaf midribs and the junction between the adaxial petiole and stem (Figure 1C, G; Figure S1) and trichome production resumes in the inflorescence (Figure 1B, E). For convenience, we refer to this restricted trichome phenotype as bald and that of *Kickxiella* species as hairy. Two aspects of trichome formation are therefore related to developmental phase: morphology changes after m4 in *Kickxiella* species while leaf blades and stems become bald from this point in subsections *Antirrhinum* and *Streptosepalum*. In a survey of 192 wild populations, we found only four that were polymorphic for bald and hairy phenotypes, possibly as a result of hybridisation with neighbouring species.

To investigate the genetic basis for variation in trichome distribution, we first crossed hairy *A. charidemi* (SE Spain *Kickxiella*; Figure 1H, J; Figure S1) with bald *A. majus* (subsection *Antirrhinum*, Figure 1B-G; Figure S1). F1 progeny had the bald phenotype of *A. majus*, suggesting that this parent carries a dominant inhibitor of trichome fate. Hairy plants occurred in the F2 at a frequency of ~6% (18 among 285) and in a genome scan were all found to be homozygous for the *A. charidemi CYC* allele, linked to the self-incompatibility (*S*) locus that prevents an active *S* allele, as from *A. charidemi* (S^c), becoming homozygous [28]. *A. majus*, in contrast, carries an inactive *s* allele [29]. Therefore the inheritance of baldness can be explained by a single locus, *Hairy* (*H*), at which the *A. majus* allele, H^m, is a dominant suppressor of trichomes and hairy F2 plants (h^c/h^c homozygotes) occur only when h^c is uncoupled from S^c by recombination.

We then screened a population of near-isogenic lines (NILs), produced from the F1 hybrid by repeated back-crossing to *A. majus*, to identify a NIL that remained heterozygous at *H* but did not carry S^c. It produced ~25% hairy progeny after self-pollination, supporting the view that the single repressor locus, *H*, is responsible for differences in trichome distribution between the parents. The NILs revealed three further aspects of H function. 1) Hairy progeny made glandular trichomes throughout vegetative development, rather than the short glandless trichomes of *A. charidemi*, implying that genes other than *H* are responsible for differences in trichome morphology (Figure S1T-W). 2) bald and hairy siblings did not differ significantly in the frequency of trichomes produced from the adaxial midrib (Table S1), which have a similar morphology in all species and are found in all leaves (Figure S1). This suggests that *H* is not active in the adaxial midrib and absence of its activity can explain the presence of midrib trichomes. Thirdly, bald and hairy did not differ significantly for two other traits related to developmental phase; flowering time and the transition from decussate to spiral phyllotaxy, which is controlled independently of flowering time [30] (Table S1). This suggests that *H* responds to developmental phase, but does not contribute to the underlying phase information.

The *H* locus was further mapped by sequencing DNA pools from hairy or bald offspring of the NIL. Mapping reads onto a draft of the parental *A. majus* genome showed the introgression from *A. charidemi* in the NIL extended over ~4.4 Mb and that *H* was located within a region of ~104 kb in which only *A. charidemi* SNPs were detected in the hairy (h^c/h^c) pool (Figure 2A, B).

RNA-seq of bald and hairy NILs identified four genes within the target region that were expressed in vegetative apices. *GRX8b* encoded a glutaredoxin in the land plant-specific CC-clade (Figure S2) and its RNA was present in bald apices but undetectable in hairy, suggesting that *H* encoded the GRX and that h^c is a null mutation (Figure 2C). The *A. charidemi* allele carried a frame-shift mutation in the *GRX8b* coding sequence resulting in nonsense mutations and promoter insertions or deletions (Figure S3) that could explain lack of transcript accumulation. A similar sequence (*GRX8a*) was found in the target region by homology (Figure 2B), but was unlikely to contribute to H function because its expression could not be detected in any aerial tissue (Figure S4). The other genes in the interval showed similar levels of expression in bald and hairy phenotypes (Figure 2C, Table S2).

If *GRX8b* is *H*, then reducing its activity by virus-induced gene silencing (ViGS) [31] should cause a hairy phenotype because *H* is a dominant suppressor of trichomes. *A. majus* infected with viruses carrying part of the *Phytoene Desaturase* (*PDS*) gene reported ViGS as tissue bleaching (Figure 3A). Whereas reducing PDS activity alone had no effect on trichome production (Figure 3C), adding either the coding region or the 3'UTR of the candidate *GRX8b* to the virus allowed trichomes to form from bleached leaves and stems above m4 (Figure 3D-F), supporting the view that *GRX8b* is needed for H activity. ViGS appeared to be specific for the *H*-linked *GRX8b*, because the most similar expressed paralogue, *GRX6c*, was unaffected (Figure 3G). Therefore *H* is very likely to be *GRX8b*. In contrast to *H*, ViGS with *GRX6c* did not alter trichome development, suggesting that this gene has a different role (Figure S5). Ectopic trichomes only formed from bleached regions of leaves infected with virus carrying *PDS* and *H* sequences, though not all bleached areas formed ectopic trichomes (Figure 3E, F). This might reflect a difference in the time or location at which *PDS* and *H* activity are required or were silenced, or that absence of *H* expression does not allow trichome formation from all parts of the leaf epidermis.

H is needed to suppress trichome formation only between m4 and the inflorescence. Consistent with this spatially restricted activity, we detected expression of the H^m allele only in apices producing bald leaves and internodes (Figure 4A). Within these, in situ hybridisation showed *H* RNA in the adaxial and abaxial epidermal cells of leaf primordia from around the developmental stage at which trichomes begin to form in hairy plants (Figure 4B, C). Expression was not detected in epidermal cells of the adaxial midrib, which produce trichomes in all genotypes (Figure 4B). Therefore the pattern of *H* activity and the distribution of trichomes appear to reflect *H* RNA expression. In contrast to *H*, mRNA from the *GRX6c* paralog was present in all aerial tissues tested and restricted to internal cells (Figure 4E-H),

supporting a role that is not related to trichome development. However, like *H*, its expression changed with developmental phase (Figure 4E), suggesting that their common ancestor was phase-regulated.

To understand the origin of H activity, we identified the sequences most similar to *H* from two other members of the tribe *Antirrhineae* (*Misopates orontium* and *Chaenorrhinum origanifolium*) and in genome sequences from more distant eudicot lineages [32, 33]. Phylogenetic analysis placed *H* in a well-supported clade specific to the *Antirrhineae*, and clustered *Antirrhinum H* with the *M. orontium* and *C. origanifolium* sequences, implying their orthology (Figure 5A). To test further the function of the *M. orontium* orthologue (*MoHairy*) we reduced its expression by ViGS. Uninfected plants become almost bald at higher vegetative metamers (Figure 6B) whereas stems experiencing ViGS were able to produce dense trichomes throughout development (Figure 6C-E), revealing a conserved role in trichome suppression for *H* and *MoHairy*. The more basal *Antirrhinum* genes in the same clade do not appear to regulate trichomes--*GRX6c* is expressed internally and ViGS with this gene does not affect trichome development, while *GRX8a* is not expressed in shoots. Therefore *H* seems likely to have acquired its trichome repressing role relatively late in the evolutionary history of angiosperms, within the Lamiales after divergence of the lineages leading to Mimulus and the *Antirrhineae* but before the *Antirrhinum*–*Misopates* split. This is consistent with H being co-opted to repress a more ancient mechanism that promotes multicellular trichome fate.

We then examined whether *H* could account for variation in trichome distribution throughout the genus *Antirrhinum*. Except for rare polymorphic populations, the hairy phenotype is unique to all alpine *Kickxiella* species [26, 27]. However these species have evolved similar complex phenotypes in parallel, raising the question of whether the same genes or mutations are responsible for parallel evolution of their hairy phenotypes. Similarly, the lowland species of subsection *Streptosepalum* arose within the basal *Kickxiella* lineage and so are likely to have evolved their restricted trichome distribution in parallel to subsection *Antirrhinum* [27] (Figure 1A).

To test these hypotheses, we first crossed the heterozygous h^c/H^m NIL to different species. Each cross to a hairy *Kickxiella* species, or to a hairy member of a polymorphic population, produced ~50% hairy progeny, suggesting that all hairy taxa lacked H activity (Table S3). Conversely all progeny of crosses to bald species, with one exception, were bald, consistent with active *H* alleles in bald species. Therefore H activity accounted for almost all the variation in trichome pattern within the genus. The exception, *A. siculum*, produced only bald offspring when crossed with *A. majus* (*H*/*H*). Though crossing to the hairy NIL (*h*/*h*) produced hairy F1 progeny, these gave rise to hairy and bald F2 phenotypes in a ~3:1 ratio (476:132, χ^2 p=0.06), suggesting that *A. siculum* is an *h* mutant but is bald because it also lacks activity of a gene required for trichome formation. This second-site mutation could have been involved in an evolutionary reversal from the hairy to bald state.

We then related trichome patterns to *H* haplotypes. The whole locus appeared to have been deleted from several closely related members of the basal *Kickxiella* lineage (Table S4, Figure S6B, C). Sequences from the remaining *Kickxiella* species, with the exception of *A. grossi*, formed a single clade, consistent with a single *h* loss-of-function mutation at its base (Figure 6). None of the alleles tested from this clade produced detectable RNA (open circles in Figure 6), suggesting that the ancestral mutation abolished expression and that additional substitutions and deletions accumulated in the absence of purifying selection. None of the coding region polymorphisms shared by this clade seem likely to prevent transcript accumulation, suggesting that the causal mutation may be in the promoter region, which is repeat-rich and highly polymorphic between species (Figure S3B). All *h* haplotypes from the later-evolved SE Spain *Kickxiella* species also belonged to this clade (underlined in Figure 6), implying that the *h* mutation involved in the early divergence of the basal *Kickxiella* lineage contributed to later parallel evolution of alpine morphology. The clade also contained non-functional *h* alleles from rare polymorphic populations of *A. latifolium*, *A. australe* and *A. graniticum*, supporting their transfer from the *Kickxiella* lineages by hybridisation. These three species are self-incompatible, and frequency-dependent selection favouring rare *Kickxiella S* alleles might have helped maintain linked *h* alleles after recent hybridisation [34]. A single *Kickxiella* haplotype was also found in all *A. siculum* samples, corroborating genetic evidence for *A. siculum* being an *h* mutant.

The active *H* alleles from bald species in subsections *Antirrhinum* and *Streptosepalum* formed a separate clade, sister to the *Kickxiella* alleles, with relatively little diversity and extensive sharing of haplotypes between species (Figure 6; Figure S6B). Therefore the *Streptosepalum* species, which evolved their lowland morphology in parallel to subsection *Antirrhinum*, appears to carry the same functional *H* alleles as subsection *Antirrhinum*. However, this clade also included four inactive *h* alleles, as inferred from allelism tests, each differing from its ancestral active allele by at least one non-synonymous substitution that could explain loss-of-function. They included *h* alleles from polymorphic populations of *A. australe* and *A. meonanthum*, which share an amino acid substitution (F78L), consistent with an independent mutation having contributed to trichome polymorphism in two populations of these species. They also included an *h* allele from *A. grosii*, a member of the basal *Kickxiella* lineage, suggesting that an older *h* mutation in this lineage had been replaced though hybridisation. The hairy character therefore appears to have evolved in parallel several times (Figure 6) and possibly reverted once to baldness in the ancestor of *A. siculum*, via a second-site suppressor mutation. However, the later parallel evolution of *Kickxiella* morphology in southeast Spain and of lowland morphology in subsection *Streptosepalum* appears to have used the same alleles that were involved in the earliest divergence of the basal *Kickxiella* and *Antirrhinum* lineages.

Trichomes are associated with beneficial roles, raising the questions of how the bald character and H function arose. Similarly, the reversal from bald to hairy appears relatively unconstrained because ViGS

revealed that it requires only loss of H activity, yet many species remain bald. To investigate whether *H* has pleiotropic effects that might constrain its evolution, we compared development of bald and hairy progeny of the NIL. Hairy (h^c/h^c) plants produced significantly larger leaves from m6 onwards compared to their bald (H^m/H^m) siblings (Figure 7). Although a gene closely linked to *H* could explain this difference, it was detected only in leaves that express *H*. This supports a pleiotropic effect of *H* in suppressing both trichome formation and leaf growth that might expose the locus to selection acting in different ways. Because variation in leaf shape and size between metamers (heteroblasty) is related to developmental phase [35], this also suggests that *H* might be involved in the response of both heteroblasty and trichome suppression to underlying developmental phase information.

DISCUSSION

H encodes a member of the land plant-specific CC clade of GRXs. Other members of this clade act as adaptors between DNA-binding TGA transcription factors and transcriptional co-repressors of the TOPLESS (TPL) family [36]. As potential oxioreductases they may also catalyse reversible redox modification of target proteins [37, 38]. Within the CC clade, the parologous ROXY1 (ROX1) and ROX2 proteins of Arabidopsis provide a precedent for developmental regulation. ROX1 binds to the TGA PERIANTHIA, reducing its ability to promote petal primordia [37, 39, 40] and acts with ROX2 to repress somatic cell fate during microspore development [38], a function that predates the monocot-dicot divergence [41, 42]. H shares the motifs required in other GRXs for oxioreductase activity and in ROX1 for interaction with TGAs and TPL co-repressors (Figure S2A) and therefore might inhibit TGAs or other factors that promote trichome fate. Although no TGA is known to be involved in trichome development, identification of H-interacting proteins will allow this to be tested.

H function appears confined to part of the Lamiales, while multicellular trichomes similar in morphology to those of *Antirrhinum* are more widespread, suggesting that *H* was recruited to control an existing mechanism of trichome development. This idea is also consistent with the evidence that *H* is not necessary for the normal spacing or morphology of trichomes. In terms of its recruitment, *H* has diverged in function from its closest expressed paralogue in *Antirrhinum*, *GRX6c*, which is phase-regulated but not expressed in the epidermis. Several other CC-clade GRX proteins appear functionally interchangeable within species so their distinct roles likely reflect their different domains of expression [36, 41, 43]. A shift in *GRX* expression to the epidermis, following gene duplication in the Lamiales, may therefore have been sufficient to bring trichome specification under control of *H*.

Evolutionary changes in trichome pattern within the genus *Antirrhinum* appear relatively unconstrained by either mutation or gene-flow. Yet, with the exception of a few polymorphic populations, trichome pattern has remained correlated with the many other morphological characters that define subsections [27] and reflect differences at multiple, unlinked loci [35, 44, 45]. This is true even of *A. grossi*, which

appears to carry a recent *h* mutation, and *A. siculum* in subsection *Antirrhinum*, which carries a second-site suppressor of its mutant *h* allele. Such persistence and parallel evolution of character combinations suggests that multiple characters, which include trichome pattern, are kept together by divergent selection. Though *H* may also suppress leaf growth, selection on leaf size is unlikely to have driven variation in trichomes because *Kicxkiella* species have the smallest leaves in the genus [27, 30]. A more plausible possibility is that a fitness cost of producing or bearing trichomes favours the bald character in the lowland habitats to which subsections *Antirrhinum* and *Streptosepalum* are adapted or in the large, upright morphologies of these subsections. Like bald *Antirrhinum* species, many angiosperms have a restricted distribution of trichomes, often related to developmental phase (reviewed by [46, 47]), consistent with a benefit from limiting them to the most vulnerable or costly parts of the plant. NILs differing only for H function could be used to examine these possibilities for *Antirrhinum*.

Parallel evolution of alpine *Kickxiella* morphologies in SE Spain and of lowland morphologies in subsection *Streptosepalum* appears to have used the same *H* alleles that were involved in the early divergence of subsection *Antirrhinum* and the basal *Kickxiella* lineage. Reuse of alleles in parallel evolution could be explained either by selection of unsorted ancestral polymorphisms or by hybridisation (reviewed by [48]). Although we identified *H* polymorphisms in four populations of subsections *Antirrhinum* and *Streptosepalum*, their presence can be explained more simply by recent hybridisation than by incomplete lineage sorting because the polymorphisms were localised and involved *h* alleles found in nearby *Kickxiella* species. However, parallel studies of the genes underlying the other characters that define morphological subsections and of their genetic backgrounds should further distinguish between these possibilities.

ACKNOWLEDGEMENTS

YT and MB were supported by studentships from Darwin Trust of Edinburgh and the Biotechnology and Biological Sciences Research Council (BBSRC; grant BB/J01446X/1), respectively. We thank Yongbiao Xue for the *A. majus* reference sequence before publication, Annabel Whibley and Enrico Coen for *A. sempervirens* and *A. pseudomajus* resequencing data, Yana Aleksandrova for initial mapping studies, Fei Yue and Douglas Pyott for advice on ViGS, Des Bradley for advice on in situ hybridisation and Jill Harrison and Justin Goodrich for comments on the manuscript.

AUTHOR CONTRIBUTIONS

YT, MB, AM and AH designed experiments and analysed data. YT, MB, YW, AB and AH conducted experiments. AH wrote the paper with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1 Variation in *Antirrhinum* trichome distribution and morphology

A) Division of the genus *Antirrhinum* into two main lineages and the distribution of alpine and lowland species between them. B-G) *A. majus* (subsection *Antirrhinum*); B) inflorescence, C) m5 stem and node, D) m2 stem, E) all aerial parts, F) short glandular trichomes on m2 leaf blades and G) the adaxial midrib of an m5 leaf. H-K) m5 leaves of two *Kickxiella* species: *A. charidemi* (SE Spain *Kickxiella*, G, I) with short glandless trichomes and *A. lopesianum* (Basal *Kickxiella*; H, J) with long glandular trichomes. Scale bars 0.1 mm. Trichome phenotypes are shown in more detail and for more species in Figure S1.

Figure 2. Mapping the *Hairy* locus

A) A scan of F_{ST} between pools of hairy or bald progeny from an h^c/H^m near-isogenic line, shown in 10 kb bins with sequence scaffold order inferred from recombination maps. Around two-thirds of plants in the bald pool will be h^c/H^m heterozygotes, giving an expected F_{ST} of ~0.5 for a marker at *H*. B) The closest recombination breakpoints in hairy (h^c/h^c) progeny delimit *H* to a region contained in three sequence scaffolds. Orange boxes show genes; the relative orientation of Scaffold 1097 could not be determined. C) Mapping of RNA-seq reads from bald and hairy phenotypes, with transcripts per million reads (TPM) values shown for the bald phenotype. Genes 1 and 5 produced no detectable transcripts and were identified by homology. Further details of all genes are given in Table S2 and Figures S2-S4.

Figure 3 *Hairy*-linked GRX8b activity is required to suppress trichome formation

A) Tissues bleaching in *A. majus* following ViGS of *Phytoene desaturase* (*PDS*) with TRV carrying *PDS* and *Hairy* sequences. B-D) ViGS with part of the 3'-UTR of *Hairy* and *PDS* allows trichome development from bleached areas of stems and leaves at metamer 5 and above (D-F), whereas silencing *PDS* alone (B) has no effect on trichomes. E-F) Trichome formation from the leaf blade is limited to regions showing ViGS—the same leaf is shown in transverse section at the position of the broken line in (E) and in a scanning electron micrograph in (F). Two spherical objects on the right-hand part of the leaf blade in F) are bubbles. G) Expression of the *Hairy*-linked *GRX8b* gene (*H*), but not its paralogue *GRX6c* (*G6c*), is reduced by virus carrying either the *Hairy*-linked *GRX8b* coding region (*cH*) or it 3'-UTR (*3'H*), while *Hairy-GRX* expression is not reduced significantly by ViGS of *GRX6c*. Each value is the mean (±SEM) from three different plants and differences with $p \leq 0.01$ are shown with different letters. Sequences used for ViGS and additional phenotypes are shown in Figure S5.

Figure 4 *Hairy* is expressed in the epidermis of bald leaves

A) Quantitative RT-PCR of *Hairy* mRNA, relative to *Ubiquitin* (*Ubi*), in vegetative apices at different stages of development. B-C) In situ detection of *Hairy* RNA in transverse sections of bald (H^m/H^m) or hairy (h^c/h^c) apices. B) *H* expression is detected in the adaxial (ad) leaf epidermis, except over the adaxial midrib, from which trichomes (arrowheads) form, and more strongly in the abaxial (ab)

epidermis. The section passes through the petiole (Pet) region of an older primordium, in which the midrib makes up more of the width than in the distal part of the leaf (Lf). E) mRNA levels of the *Hairy* paralogue *GRX6c* and F-H) in situ hybridisation with a *GRX6c* probe. Tests for expression of the *Hairy*-linked *GRX8a* gene are shown in Figure S3

Figure 5 Origin of the *Hairy* gene

A) Maximum-likelihood tree of *A. majus* (Am) Hairy protein and the most similar sequences from *A. majus* (red), *Misopates orontium* (MoHairy), *Chaenorrhinum origanifolium* (CoGRX), Mimulus (orange), tomato (green) and *A. thaliana* (blue). *A. majus GRX* genes are numbered according to their chromosome location. Bootstrap support ≥50% is shown. The tree is rooted on a sister clade (Figure S6A). B) *M. orontium* stems become bald from metamer 3. C) ViGS of *MoHairy* and *PDS* genes. The right-hand part of the stem shows silencing, revealed by tissue bleaching in a transverse section (D), with ectopic trichomes shown in a negative image of the same section (arrowheads in E).

Figure 6. Evolution of the *Hairy* gene within *Antirrhinum*

A maximum-likelihood tree of the *H* open reading frame. Functional *H* alleles, inferred from phenotypes and allelism tests, are shown in red and non-functional *h* alleles in blue. Underlined alleles are from the SE Spain *Kickxiella* species that evolved alpine morphology later in parallel. Expressed alleles are marked with filled circles, those for which no expression could be detected with open circles. Nodes recovered in ≥50% of bootstrap replicates are marked with dots. More than one allele was identified in many species; the distribution of haplotypes between species and populations is given in Table S4 and Figure S6. Non-synonymous substitutions that can explain independent losses of function are in grey boxes. The *A. australe* 3 allele shares the F78L substitution with two other non-functional *h* haplotypes, but also has a unique deletion of eight amino acids, shown by an artificially lengthened terminal branch that is not to scale.

Figure 7 The *H* locus affects leaf area

A) A comparison of leaf areas for hairy (h^c/h^c, blue) and bald (H^m/H^m red) siblings in the same genetic background. Values are means ±SE for 13 hairy and 16 bald plants; means significantly different at α=0.01 or 0.05 are shown by * or **, respectively. B) Mean outlines of the leaves for each genotype are shown in green, with +1 SE in grey.

STAR Methods

Key Resources Table:

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Viral Strains		
E. coli strain DH5α	Widely distributed	n/a
E. coli strain DH10B	Widely distributed	n/a
Rhizobium radiobacter strain GV3101 (pMP90RK)	Widely distributed	n/a
Antibodies		
Anti-Digoxigenin-AP Fab fragments	Roche	# 11093274910
Chemicals, Peptides, and Recombinant Proteins		
Q5 High-fidelity DNA Polymerase	New England Biolabs	# M0491
DNA restriction and modification enzymes	New England Biolabs	various
Taq DNA polymerase	[49]	n/a
M-MLV reverse transcriptase	Promega	# M1701
DNAse I (RNase free)	New England Biolabs	# M0303S
Nucleospin Gel and PCR Clean-up kit	Machery Nagel	# 40609.10
Nucleospin Plasmid kit	Machery Nagel	# 740499.240
TRIzol Reagent	Life Technologies	# 15596-026
PureLink RNA Mini Kit	Life Technologies	# 12183018A
Araldite Ultra Standard epoxy adhesive	widely available	# ARA-400001
Extrude Type 3 polyvinylsiloxane impression material	Kerr Corporation	# 28418
Safranin O	Sigma-Aldrich	# S2255
Calcofluor White/FB28	Sigma-Aldrich	# F3543
Digoxigenin-11-UTP	Roche	# 03359247910
Nitro Blue Tetrazolium Cl	Fisher Bioreagents	BP108-1
BCIP-T	Thermo Scientific	R0821
Phytagel	Sigma-Aldrich	P8169
Critical Commercial Assays		
LightCycler 480 SYBR Green I Master	Roche	# 04 707 516 001
Deposited Data		
RNAseq reads	ENA	PRJEB29128
PoolSeq reads		
H DNA sequence alignments and trees	TreeBase	Please see temporary link below
H-like protein sequence alignments and trees		
Experimental Models: Organisms/Strains		
Nicotiana benthamiana	widely available	n/a
Antirrhinum and related species	Please see Table S4	n/a
Oligonucleotides--Please see Table S6		
Recombinant DNA		
pTRV1	[31]	n/a
pTRV2	[31]	n/a
pJet1.2	Life Technologies	K1231
Software and Algorithms		
MEGA	[50]	n/a
BWA-MEM	[51]	n/a
Stampy	[52]	n/a
Trimmomatic	[53]	n/a
PoPoolation 2	[54]	n/a

TopHat/Cufflinks	[55]	n/a
IGV	[56]	n/a
MUSCLE	[57]	n/a
ProtTest	[58]	n/a
jModelTest 2	[59]	n/a
RAxML v8	[60]	n/a
AutoStitch	[61]	n/a
Fiji	[62]	n/a
AAMToolbox	http://lemur.cmp.uea.ac.uk/Research/cbg/Documents/Bangham-Coen-Group/AAMToolbox	
Other		
Levington John Innes No. 2 compost	widely available	SKU 000041836
Osmocote Pro (18-9-10) fertilizer granules	no longer available	n/a

Lead Contact and Materials Availability

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Andrew Hudson (andrew.hudson@ed.ac.uk). Recombinant DNA and plant genotypes generated in this study are available on request.

Experimental Models and Subject Details

The origins of the *Antirrhinum* species and their relatives used in this study are detailed in Table S4 and their taxonomy in [27]. Generation of a population of near-isogenic lines (NILs) has been described elsewhere [35]. A NIL carrying the *A. charidemi* allele of the *H*-linked *CYCLOIDEA* (*CYC*) locus were identified by a *Kpn* I CAPS after amplification with primers CYC-F and CYC-R (Table S6). Self-pollination of one of these NILs produced a low proportion (6%) of progeny with the hairy phenotype, suggesting that it was an H^m/h^c heterozygote and also heterozygous s^m/S^c at the self-incompatibility (*S*) locus. To obtain an H^m/h^c NIL lacking the active S^c allele, a hairy offspring, (likely genotype h^cs^m/h^cS^c), was back-crossed to JI.7 and an individual inheriting the recombinant h^cs^m haplotype identified by a *Dpn* II CAPS in an F-box component of *S* [63], amplified with primers SLF-F and SLF-R (Table S6). This haplotype had undergone a second recombination, on the opposite side of h^c from *S*, uncoupling h^c from CYC^c. The recurrent *A. majus* parent used in the generation of NILs (the inbred line JI.7) was the source of the reference *A. majus* genome sequence [65]. For allelism tests, different species were used to pollinate a heterozygous h^c/H^m NIL and the proportion of hairy and bald F1 progeny recorded (Table S3).

Almost all plants used for comparison of phenotypes or gene expression were grown in John Innes No. 2 compost supplemented with 10 g l^{-1} Osmacote slow-release fertilizer in 550 ml pots. This included plants used to test for effects of *H* on other aspects of developmental phase (Table S1). They were kept in a glasshouse at a day-time temperature of 21.5˚C (± 1.0˚C s.d.) at 62% ± 6% relative humidity and a night-time temperature of 20˚C ± 0.2˚C and 62% ± 8% relative humidity. Plants received supplemental lighting from metal halide lamps to maintain a minimum of 480 µmol m^{-2} s^{-1} PAR during a 16 hour day. The exceptions were the additional species shown in Figure S1S and plants used only for genetics or

genotype analysis, which were grown together in a heated and lit glasshouse with less stringent environmental control.

Method Details

Whole-mount preparation of leaves for light microscopy

To examine trichome morphology, leaves were vacuum infiltrated with 100% ethanol and stored until chlorophyll had been removed then rehydrated in phosphate buffered saline (PBS, 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.9 mM KH_2PO_4, pH 7.4) and softened and cleared in 0.5 M NaOH for 1 hour at 60°C. After washing in PBS, samples were stained in 1% safranin in PBS for 2 minutes and mounted in water. To examine sections of fresh leaves or stems, tissues were embedded in 7% Phytagel and ~75 µm sections cut with a vibratome and mounted in water. For sections that were wider than the microscope's field of view, multiple images were merged with AutoStitch software [61].

Scanning electron microscopy (SEM)

Resin replicas of plant surfaces were imaged by SEM. A mould was first made by spreading polyvinyl siloxane dental impression medium (Kerr Reflect) over the plant surface. When polymerised, the mould was removed from the tissue, filled with epoxy adhesive (Araldite) and subjected to a partial vacuum (-85 kPa) for 5 minutes at 60°C to remove air bubbles. The epoxy was then allowed to polymerise for 24 hours at room temperature before the replicas were removed from their moulds, mounted on SEM stubs, sputter-coated with a ~25 nm film of gold and palladium, and imaged at ambient temperature. Image stitching was used for large objects [61].

Pool-seq

DNA was extracted from individual progeny of a heterozygous NIL (H^m/h^c). For each phenotype--hairy (h^c/h^c genotype) or bald (H^c/h^c or H^c/H^c)–an equal quantity of DNA from each of 86 plants was pooled. Illumina Hi-seq libraries were prepared at Earlham Institute, Norwich and sequenced to generate 100 bp paired-end reads. Reads were cleaned with Trimmomatic [53], and aligned to a draft *A. majus* genome with Stampy [52]. A pileup file for the two pools was created from their BAM files and converted to sync format in Popoolation2 [54]. Popoolation2 was used to calculate mean F_{ST} values between the two phenotype pools for SNPs and short indels in 10 kb bins and to test significance of association between of SNPs and trichome phenotypes. Because the hairy pool carries only the h^c allele, while an average of one third of the alleles in bald plants will be h^c, a SNP at *H* has an expected F_{ST} of ~0.5. The order of genome sequence scaffolds along chromosomes was inferred from recombination mapping data.

The region containing *h* was delimited further by flanking recombination events, detected as *A. majus* polymorphisms in the pool of hairy phenotypes (genotype h^c/h^c). To minimise the possibility of falsely calling polymorphisms that were sequencing errors, the criterion set for detection of an *A. majus* SNP

was either a single sequence feature covered by at least two independent reads or the outermost of two features within 250 bp covered by different reads. Mapping of reads to the reference was also checked visually in IGV [56] to confirm that the polymorphisms flanking *h* were not the result of inconsistent read alignment around indels.

RNA-seq

RNA was extracted from vegetative shoot apices containing ≥m5 leaf primordia taken from multiple individuals of either the *hairy* NIL or its *A. majus* progenitor. RNA extraction involved grinding 200 mg of tissue in 2 ml TRIzol reagent, removal of insoluble material by centrifugation at 12,000 x*g* for 10 minutes followed by extraction with 0.4 ml of chloroform. Ethanol was added to the aqueous phase to 35% (v/v) and the mixture passed through a Purelink RNA column. Bound RNA was washed and eluted from the column according to the supplier's instructions. The vegetative state of the harvested apices was confirmed by reverse-transcription and PCR of the *FLORICAULA* gene, which is expressed only after the transition to reproductive development [65], using primers FLO-F and FLO-R (Table S6). cDNA libraries for RNA-seq were made from the RNA samples and sequenced to generate 100 bp single-end reads at Glasgow Polyomics. Reads were cleaned and trimmed as for genomic sequences, mapped onto the draft *A. majus* genome and quantified with Tuxedo suite software [55].

Virus-induced gene silencing (ViGS)

Plasmids pTRV1 and pTRV2Δ2b, carrying sequences that express modified versions of the bi-partite genome of Tobacco Rattle Virus (TRV) [31] were used for ViGS. To develop a reporter for silencing, 344 bp of the single-copy *Phytoene Desaturase* (*PDS*) gene of *A. majus* was identified by homology in the draft *A. majus* genome sequence, amplified from cDNA and cloned between *Asc* I and *Bam* HI sites in the pTRV2 vector. *PDS* and test gene sequences were fused by overlap PCR before cloning at the same position. Either the whole ORF or 237 bp of the 3'-UTR from 5 nt after the stop codon were used for *H* while the whole ORF was used for *GRX6c* (Figure S5A, B).

Agrobacterium cells (GV3101 pMP90) carrying either pTRV1 or recombinant pTRV2Δ2b were mixed and infiltrated into leaves of *Nicotiana benthamiana*. After 5-7 days, systemically infected leaves were ground in twice their mass of 1 mM phosphate buffer (pH 7.4) to obtain an infectious virus extract. Cotyledons and m2 leaves of *Antirrhinum* or *Misopates orontium* seedlings were inoculated by rubbing with the virus extract and Al_2O_3 abrasive. The effect of silencing *PDS* (bleaching) was visible in leaves and stems within ~7 days of infection. The *A. majus PDS* sequence caused ViGS in *M. orontium* so was also used to monitor silencing in this species.

***In situ* hybridization to RNA**

RNA was detected in sections of fixed tissue using the method of [64]. To produce digoxigenin-labelled RNA probes, the 3'-UTR of *A. majus H* or *GRX6c* were cloned in pJET1.2, amplified as a fusion with the vector's T7 promoter sequence and the PCR products transcribed with T7 RNA polymerase using

digoxigenin-labelled dUTP and unlabelled dATP, dCTP and dGTP. The RNA probe was not hydrolysed and was hybridised to 8 µm tissue sections. It was detected using alkaline phosphatase-coupled antibodies against digoxigenin, and alkaline phosphatase activity revealed with X-Gluc and nitro blue tetrazolium (NBT) substrates. Cell walls were counterstained with calcofluor white and sections imaged in either white light, or a combination of white and UV (365 nm) light.

Phylogenetic analyses

H sequences of *Antirrhinum* species were amplified from genomic DNA with primer pairs H-L/H-R, H-L/H-R2, or H-L3/H-R3 (Table S6), using Q5 polymerase. For the outgroup species *Chaenorrhinum origanifolium* and *M. orontium*, they were amplified from cDNA of vegetative shoot apices by 5' and 3' RACE using primers GRX-CR-R, or GRX-RACE-R3 and RACE adaptor primers [67, 68] (Table S6). Direct sequencing of PCR products indicated heterozygosity in most *Antirrhinum* accessions. Therefore products of an independent PCR were cloned in pJET1.2 and clones sequenced until both haplotypes had been recovered from a heterozygous individual. PCR-generated mutations were eliminated by comparing sequence traces from clones and initial PCR products.

The *H* coding sequences were aligned with MUSCLE in MEGA [50, 57]. MEGA was also used for maximum likelihood (ML) estimates of phylogeny, with *M. orontium* as outgroup, use of sites present in ≥75% of sequences and a model with K2P substitution and gamma distribution of rates, as chosen by jModeltest2 [59]. Bootstrap percentages are from 500 resamplings.

To examine the likely origin of H, similar proteins were identified by blast searches of the inferred proteomes of *Arabidopsis thaliana*, Mimulus (*Erythranthe guttata*) and tomato (Table S5). Other *A. majus* proteins were inferred from a combination of genome and cDNA sequences and *Misopates* and *Chaenorrhinum* proteins from amplified cDNAs.

Partial amino acid sequences (corresponding to residues 11-107 of *A. majus* H) were aligned with MUSCLE. The most likely ML tree was identified in MEGA using a JTT model with gamma distributed rates, as suggested by ProtTest2.4 [58], and tested with 500 bootstrap replicates. The tree was rooted with reference to a larger tree (Figure S6A), produced using RaxML [60] under the same model.

Quantification and Statistical Analysis

Leaf allometry

To test for effects of the *hairy* introgression on leaf shape and size, bald (n=16) and hairy (n=13) progeny of the heterozygous NIL were grown together and leaves from m2-13 harvested when the first flower opened. Leaves were flattened and scanned and leaf areas calculated with Fiji [62]. Areas of leaves at each metamer were subject to Shapiro-Wilk tests for normality and compared between genotypes with 2-tailed *t*-tests with correction for multiple testing. To compare the shapes and sizes of leaves at all nodes, each leaf outline was converted to a series of points using AAMToolbox, so that each leaf was

described by 53 2D co-ordinates. For each node, the sets of points for all plants were rotated and translated to minimise variance within the dataset. Point sets for all leaves of a plant were then combined, the mean position of each point in the hairy or bald plant set used as the mean shape for that genotype, and plotted with its standard error.

Trichome density

Variation in trichome density within individuals was quantified for two bald H^m/H^m and two hairy h^c/h^c plants grown together under the same conditions (see above). For each plant, a leaf was taken from metamer m1 (cotyledons) to m9, cleared and stained with safranin. For each side of a leaf (adaxial or abaxial), the number of trichomes was counted in in five fields of 2.41 mm^2; one positioned randomly in the distal third of the leaf's length, two in the proximal third and two in the middle third, avoiding the midrib and the leaf margin. Because the transition from hairy to bald occurred part way along m4 H^m/H^m leaves, the proximal and distal halves were sampled separately. The areas of ~20 epidermal pavement cells in each field were also estimated in Fiji, and used with trichome density values to estimate a trichome index (percentage of pavement epidermal cells bearing trichomes). For each leaf surface, the 95% confidence interval of the mean index value was estimated from the frequency distribution of $1x10^5$ trichome index calculations made from bootstrap resampled trichome and cell density values. To compare trichome indices for different species, the number of adaxial trichomes and pavement cells lacking trichomes were counted in a 0.25 mm^2 field around the midpoint of an m5 leaf, excluding the midrib and leaf margin, and used to calculate trichome index for the individual, and mean trichome index and SEM for the species. The number of members of each species sampled are shown in Figure S1R. To compare the frequency of adaxial midrib trichomes between bald and hairy NILs, we counted the number of trichomes in the basal 14.4 mm of 20 mature m5-6 leaves of each genotype. The shape of the midrib and density of trichomes obscured many underlying epidermal cells, therefore frequencies are expressed as trichomes mm^{-1}. The probability that the means of the two genotypes were the same was estimated with a 2-tailed *t*-test.

Quantitative RT-PCR

Three developmental stages were sampled: aerial parts of seedlings ~16 days after germination (with m2-m3 leaf primordia), shoot apices with 5 mm-long leaves at m5, and inflorescence apices. At least three tissue samples, each a pool of different individuals, were processed separately for each stage and genotype. RNA was extracted with Trizol reagent, purified on Invitrogen Purelink columns and incubated with DNase I to remove genomic DNA contamination, as confirmed by PCR. cDNA synthesis was primed from oligo-dT and real-time amplification from cDNA templates carried out in a Roche LightCycler 96, with at least two technical replicates of each primer-template combination. Expression was quantified relative to *A. majus Ubiquitin5* [69], with amplification efficiencies for each gene

estimated from a serial dilution of pooled experimental templates. The same procedure was used to test the effects of ViGS on gene expression, using RNA samples from three different plants from each treatment. Primer sequences are listed in Table S6. The ratio of RNA abundance for the target gene to *Ubiquitin5* was calculated for each biological replicate and differences in expression between treatments or genotypes were detected with ANOVA and Tukey's post-hoc tests.

Data Availability

Raw sequence reads from pool-seq and RNA-seq experiments and *H* gene sequences from *Antirrhinum* and related species have been deposited in the European Nucleotide Archive (ENA, study accession number PRJEB29128) and sequences, alignments, parameters and trees from phylogenetic analyses in TreeBase (http://purl.org/phylo/treebase/phylows/study/TB2:S23381).

REFERENCES

1 Bickford, C.P. (2016). Ecophysiology of leaf trichomes. Functional Plant Biology *43***,** 807-814.

2 Riddick, E.W. and Simmons, A.M. (2014). Do plant trichomes cause more harm than good to predatory insects? Pest. Manag. Sci*. 70***,** 1655-1665.

3 Carlquist, S.J. (1961). *Comparative Plant Anatomy*, 1 ed. New York, Holt, Rinehart & Winston.

4 Huchelmann, A., Boutry, M. and Hachez, C. (2017). Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. Plant Physiol *175***,** 6-22.

5 Schilmiller, A.L., Last, R.L. and Pichersky, E. (2008). Harnessing plant trichome biochemistry for the production of useful compounds. Plant J. *54*, 702-711.

6 Glas, J.J., Schimmel, B.C., Alba, J.M., Escobar-Bravo, R., Schuurink, R.C. and Kant, M.R. (2012). Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. Int. J. Mol. Sci*. 13***,** 17077-17103.

7 Evans, M.M., Passas, H.J. and Poethig, R.S. (1994). Heterochronic effects of glossy15 mutations on epidermal cell identity in maize. Development *120***,** 1971-1981.

8 Telfer, A., Bollman, K.M. and Poethig, R.S. (1997). Phase change and the regulation of trichome distribution in *Arabidopsis thaliana*. Development *124*, 645-654.

9 Bloomer, R.H., Juenger, T.E. and Symonds, V.V. (2012). Natural variation in *GL1* and its effects on trichome density in *Arabidopsis thaliana*. Mol. Ecol*. 21*, 3501-3515.

10 Vernoud, V., Laigle, G., Rozier, F., Meeley, R.B., Perez, P. and Rogowsky, P.M. (2009). The HD-ZIP IV transcription factor OCL4 is necessary for trichome patterning and anther development in maize. Plant J*. 59*, 883-894.

11 Walford, S.A., Wu, Y., Llewellyn, D.J. and Dennis, E.S. (2012). Epidermal cell differentiation in cotton mediated by the homeodomain leucine zipper gene, *GhHD-1*. Plant J. *71*, 464-478.

12 Wang, Y.L., Nie, J.T., Chen, H.M., Guo, C.L., Pan, J., He, H.L., Pan, J.S. and Cai, R. (2016). Identification and mapping of *Tril*, a homeodomain-leucine zipper gene involved in multicellular trichome initiation in *Cucumis sativus*. Theor. Appl. Genet. *129*, 305-316.

13 Yan, T., Chen, M., Shen, Q., Li, L., Fu, X., Pan, Q., Tang, Y., Shi, P., Lv, Z., Jiang, W. *et al*. (2017). HOMEODOMAIN PROTEIN 1 is required for jasmonate-mediated glandular trichome initiation in *Artemisia annua*. New Phytol. *213*, 1145-1155.

14 Yang, C., Li, H., Zhang, J., Luo, Z., Gong, P., Zhang, C., Li, J., Wang, T., Zhang, Y., Lu, Y. and Ye, Z. (2011). A regulatory gene induces trichome formation and embryo lethality in tomato. Proc. Natl. Acad. Sci. U.S.A *108*, 11836-11841.

15 Zhu, H., Sun, X., Zhang, Q., Song, P., Hu, Q., Zhang, X., Li, X., Hu, J., Pan, J., Sun, S. *et al*. (2018). *GLABROUS* (*CmGL*) encodes a HD-ZIP IV transcription factor playing roles in multicellular trichome initiation in melon. Theor. Appl. Genet. *131*, 569-579.

16 Chang, J., Yu, T., Yang, Q., Li, C., Xiong, C., Gao, S., Xie, Q., Zheng, F., Li, H., Tian, Z. *et al*. (2018). *Hair*, encoding a single C2H2 zinc-finger protein, regulates multicellular trichome formation in tomatoes. Plant J. *96*, 90-102.

17 Machado, A., Wu, Y., Yang, Y., Llewellyn, D.J. and Dennis, E.S. (2009). The MYB transcription factor ChMYB25 regulates early fibre and trichome development. Plant J *59*, 52-62.

18 Shi, P., Fu, X., Shen, Q., Liu, M., Pan, Q., Tang, Y., Jiang, W., Lv, Z., Tan, T., Ma, Y. *et al*. (2018). The role of AaMIXTA1 in regulating the initiation of glandular trichomes and cuticle biosynthesis in *Artemisia annua*. New Phytol. *217*, 261-276.

19 Martin, C., Bhatt, K., Baumann, K., Jin, H., Zachgo, S., Roberts, K., Schwarz-Sommer, Z., Glover, B. and Perez-Rodrigues, M. (2002). The mechanics of cell fate determination in petals. Philos. Trans. R. Soc. Lond. B Biol. Sci. *357*, 809-13.

20 Pesch, M. and Hulskamp, M. (2009). One, two, three...models for trichome patterning in Arabidopsis? Curr. Opin. Plant Biol. *12*, 587-592.

21 Serna, L. and Martin, C. (2006). Trichomes: different regulatory networks lead to convergent structures. Trends Plant Sci. *11*, 274-80.

22 Chen, E., Zhang, X., Yang, Z., Wang, X., Yang, Z., Zhang, C., Wu, Z., Kong, D., Liu, Z., Zhao, G. *et al*. (2017). Genome-wide analysis of the HD-ZIP IV transcription factor family in *Gossypium arboreum* and *GaHDG11* involved in osmotic tolerance in transgenic Arabidopsis. Mol. Genet. Genomics *292*, 593-609.

23 Zalewski, C.S., Floyd, S.K., Furumizu, C., Sakakibara, K., Stevenson, D.W. and Bowman, J.L. (2013). Evolution of the class IV HD-zip gene family in streptophytes. Mol. Biol. Evol. *30*, 2347-2365.

24 Javelle, M., Klein-Cosson, C., Vernoud, V., Boltz, V., Maher, C., Timmermans, M., Depege-Fargeix, N. and Rogowsky, P.M. (2011). Genome-wide characterization of the HD-ZIP IV transcription factor family in maize: preferential expression in the epidermis. Plant Physiol. *157*, 790-803.

25 Rothmaler, W. (1956). Taxonomische Monographie der Gattung *Antirrhinum*. Feddes Repertorium 136, 1-124.

26 Sutton, D.A. (1988). *A revision of the tribe Antirrhineae* Oxford, Oxford University Press.

27 Wilson, Y. and Hudson, A. (2011). The evolutionary history of *Antirrhinum* suggests that ancestral phenotype combinations survived repeated hybridizations. Plant J*. 66*, 1032-1043.

28 Schwarz-Sommer, Z., de Andrade, S.E., Berndtgen, R., Lonnig, W.E., Muller, A., Nindl, I., Stuber, K., Wunder, J., Saedler, H., Gubitz, T. *et al*. (2003). A linkage map of an F2 hybrid population of *Antirrhinum majus* and *A. molle*. Genetics *163*, 699-710.

29 Qiao, H., Wang, F., Zhao, L., Zhou, J., Lai, Z., Zhang, Y., Robbins, T.P. and Xue, Y. (2004). The F-box protein AhSLF-S2 controls the pollen function of S-RNase-based self-incompatibility. Plant Cell *16*, 2307-2322.

30 Bradley, D., Vincent, C., Carpenter, R. and Coen, E.S. (1996). Pathways for inflorescence and floral induction in Antirrhinum. Development *122*, 1535-44.

31 Liu, Y., Schiff, M., Marathe, R. and Dinesh-Kumar, S.P. (2002). Tobacco *Rar1*, *EDS1* and *NPR1*/*NIM1 like* genes are required for N-mediated resistance to tobacco mosaic virus. Plant J*. 30*, 415-429.

32 Ogutcen, E. and Vamosi, J.C. (2016). A phylogenetic study of the tribe Antirrhineae: Genome duplications and long-distance dispersals from the Old World to the New World. Am. J. Bot. *103*, 1071-1081.

33 Zeng, L., Zhang, N., Zhang, Q., Endress, P.K., Huang, J. and Ma, H. (2017). Resolution of deep eudicot phylogeny and their temporal diversification using nuclear genes from transcriptomic and genomic datasets. New Phytol. *214*, 1338-1354.

34 Wright, S. (1939). The Distribution of Self-Sterility Alleles in Populations. Genetics *24*, 538-552.

35 Costa, M.M., Yang, S., Critchley, J., Feng, X., Wilson, Y., Langlade, N., Copsey, L. and Hudson, A. (2012). The genetic basis for natural variation in heteroblasty in *Antirrhinum*. New Phytol. *196*, 1251-1259.

36 Uhrig, J.F., Huang, L.J., Barghahn, S., Willmer, M., Thurow, C. and Gatz, C. (2017). CC-type glutaredoxins recruit the transcriptional co-repressor TOPLESS to TGA-dependent target promoters in *Arabidopsis thaliana*. Biochim. Biophys. Acta *1860*, 218-226.

37 Xing, S., Rosso, M.G. and Zachgo, S. (2005). ROXY1, a member of the plant glutaredoxin family, is required for petal development in *Arabidopsis thaliana*. Development *132*, 1555-1565.

38 Xing, S., Zachgo, S. (2008). *ROXY1* and *ROXY2*, two Arabidopsis glutaredoxin genes, are required for anther development. Plant J. *53*, 790-801.

39 Li, S., Lauri, A., Ziemann, M., Busch, A., Bhave, M. and Zachgo, S. (2009). Nuclear activity of ROXY1, a glutaredoxin interacting with TGA factors, is required for petal development in *Arabidopsis thaliana*. Plant Cell *21*, 429-441.

40 Murmu, J., Bush, M.J., DeLong, C., Li, S., Xu, M., Khan, M., Malcolmson, C., Fobert, P.R., Zachgo, S. and Hepworth, S.R. (2010). Arabidopsis basic leucine-zipper transcription factors TGA9 and TGA10 interact with floral glutaredoxins ROXY1 and ROXY2 and are redundantly required for anther development. Plant Physiol. *154*, 1492-1504.

41 Wang, Z., Xing, S., Birkenbihl, R.P. and Zachgo, S. (2009). Conserved functions of Arabidopsis and rice CC-type glutaredoxins in flower development and pathogen response. Mol. Plant *2*, 323-335.

42 Yang, F., Bui, H.T., Pautler, M., Llaca, V., Johnston, R., Lee, B.H., Kolbe, A., Sakai, H. and Jackson, D. (2015). A maize glutaredoxin gene, *abphyl2*, regulates shoot meristem size and phyllotaxy. Plant Cell *27*, 121-131.

43 Li, S., Gutsche, N. and Zachgo, S. (2011). The ROXY1 C-terminal L**LL motif is essential for the interaction with TGA transcription factors. Plant Physiol. *157*, 2056-2068.

44 Feng, X., Wilson, Y., Bowers, J., Kennaway, R., Bangham, A., Hannah, A., Coen, E. and Hudson, A. (2009). Evolution of allometry in *Antirrhinum*. Plant Cell *21*, 2999-3007.

45 Langlade, N.B., Feng, X., Dransfield, T., Copsey, L., Hanna, A.I., Thébaud, C., Bangham, A., Hudson, A. and Coen, E. (2005). Evolution through genetically controlled allometry space. Proc. Natl. Acad. Sci. U.S.A. *102*, 10221-10226.

46 Johnson, H.B. (1975). Plant pubescence: an ecological perspective. Botanical Review *41*, 233-258

47 Poethig, R.S. (1990). Phase change and the regulation of shoot morphogenesis in plants. Science *250*, 923-930.

48 Marques, D.A., Meier, J. and Seehausen, O. (2019). A combinatorial view on speciation and adaptive radiation. Trends Ecol. Evol. *34*, 531-44.

49 Engelke, D.R., Krikos, A., Bruck, M.E. and Ginsburg, A. (1990). Purification of *Thermus aquaticus* DNA polymerase expressed in *Escherichia coli*. Anal. Biochem. *191*, 369-400.

50 Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol. Biol. Evol. *35*, 1547-49.

51 Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, 1303.3997v2

52 Lunter, G. and Goodson, M. (2011). Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res*. 21*, 936-939.

53 Bolger, A.M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

54 Kofler, R., Pandey, R.V. and Schlötterer, C. (2011). PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **27:**3435-3436.

55 Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. *7*, 562-578.

56 Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24-26.

57 Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res*. 32*, 1792-1797.

58 Abascal, F., Zardoya, R. and Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. Bioinformatics *21*, 2104-2105.

59 Darriba, D., Taboada, G.L., Doallo, R. and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods *9*, 772.

60 Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312-1313.

61 Brown, M. and Lowe, D. (2007). Automatic panoramic image stitching using invariant features. Intl. *J.* Computer Vision *74*, 59-73.

62 Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. *et al*. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676-682.

63 Lai, Z., Ma, W., Han, B., Liang, L., Zhang, Y., Hong, G. and Xue, Y. (2002). An F-box gene linked to the self-incompatibility (*S*) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. Plant Mol. Biol. *50*, 29-42.

64 Li, M., Zhang, D., Gao, Q., Luo, Y., Zhang, H., Ma, B., Chen, C., Whibley, A., Zhang, Y., Cao, Y. *et al.* (2019). Genome structure and evolution of *Antirrhinum majus* L. Nature Plants *5*, 174–183.

65 Bradley, D., Carpenter, R., Copsey, L., Vincent, C., Rothstein, S. and Coen, E. (1996). Control of inflorescence architecture in *Antirrhinum*. Nature *379*, 791-797.

66 Rebocho, A.B., Southam, P., Kennaway, J.R., Bangham, J.A. and Coen, E. (2017). Generation of shape complexity through tissue conflict resolution. Elife *6*, e20156.

67 Scotto-Lavino, E., Du, G. and Frohman, M.A. (2006). 3' end cDNA amplification using classic RACE. Nat. Protoc. *1*, 2742-2745.

68 Scotto-Lavino, E., Du, G. and Frohman, M.A. (2006) Amplification of 5' end cDNA with 'new RACE'. Nat. Protoc. *1*, 3056-3061.

69 Preston, J.C. and Hileman, L.C. (2010). SQUAMOSA-PROMOTER BINDING PROTEIN 1 initiates flowering in *Antirrhinum majus* through the activation of meristem identity genes. Plant J. *62*, 704-712.

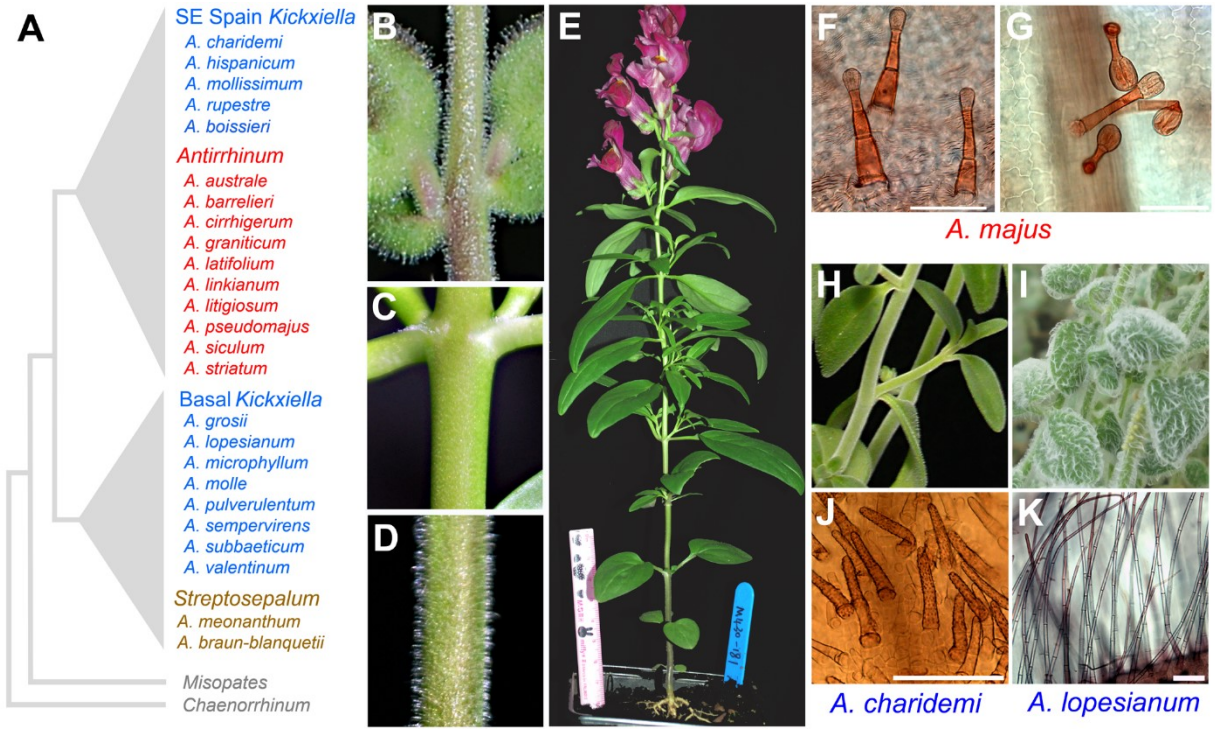


Figure 1

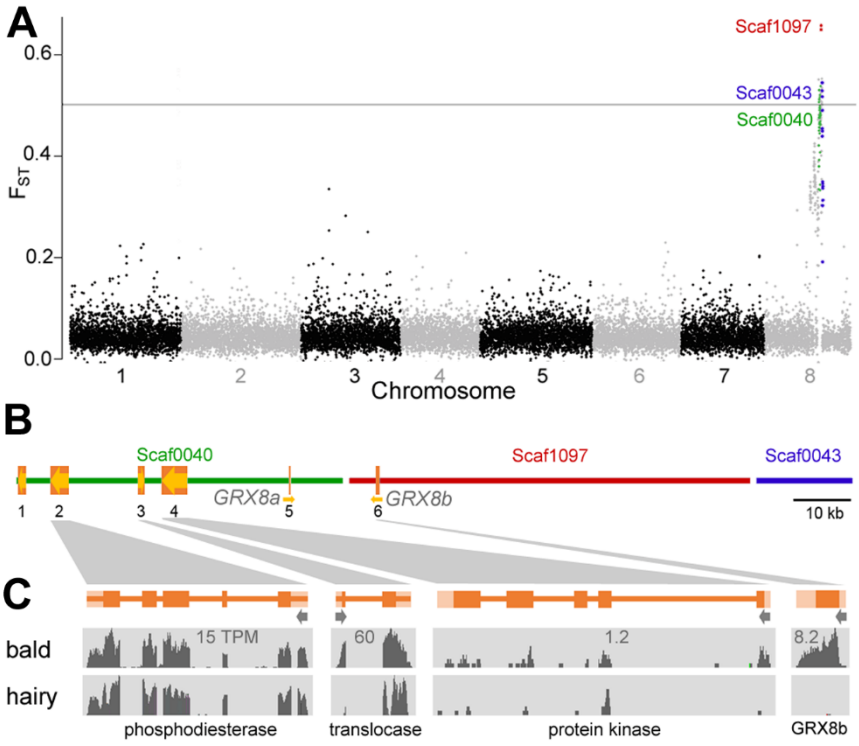


Figure 2

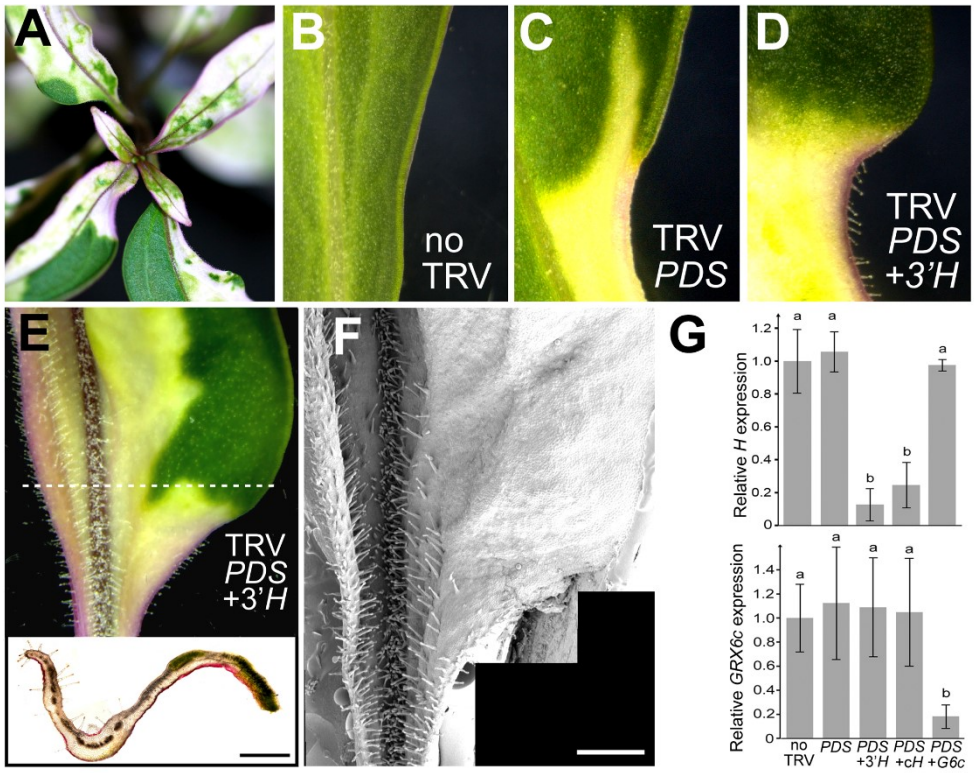


Figure 3

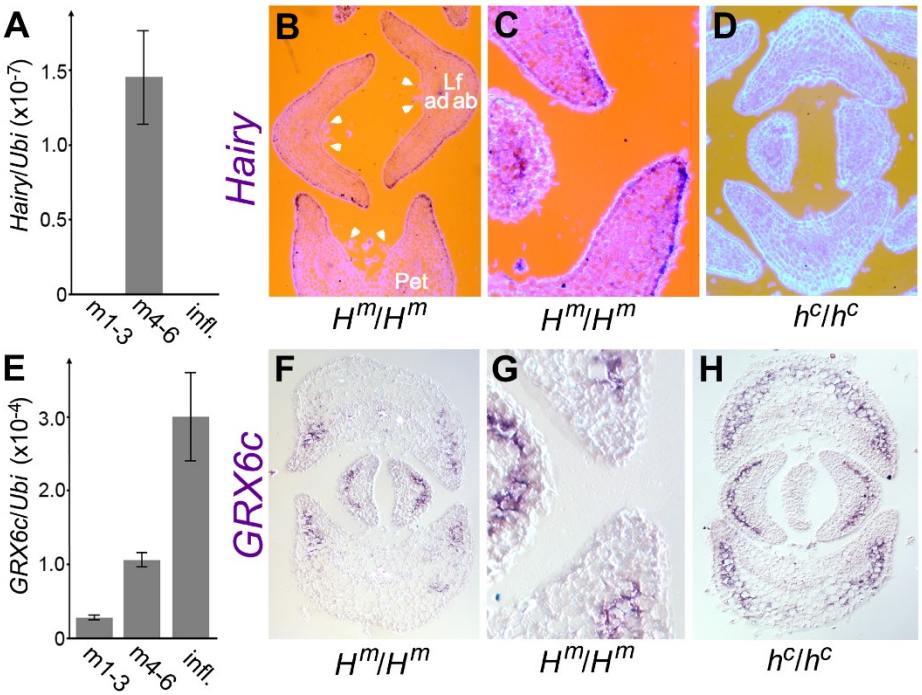


Figure 4

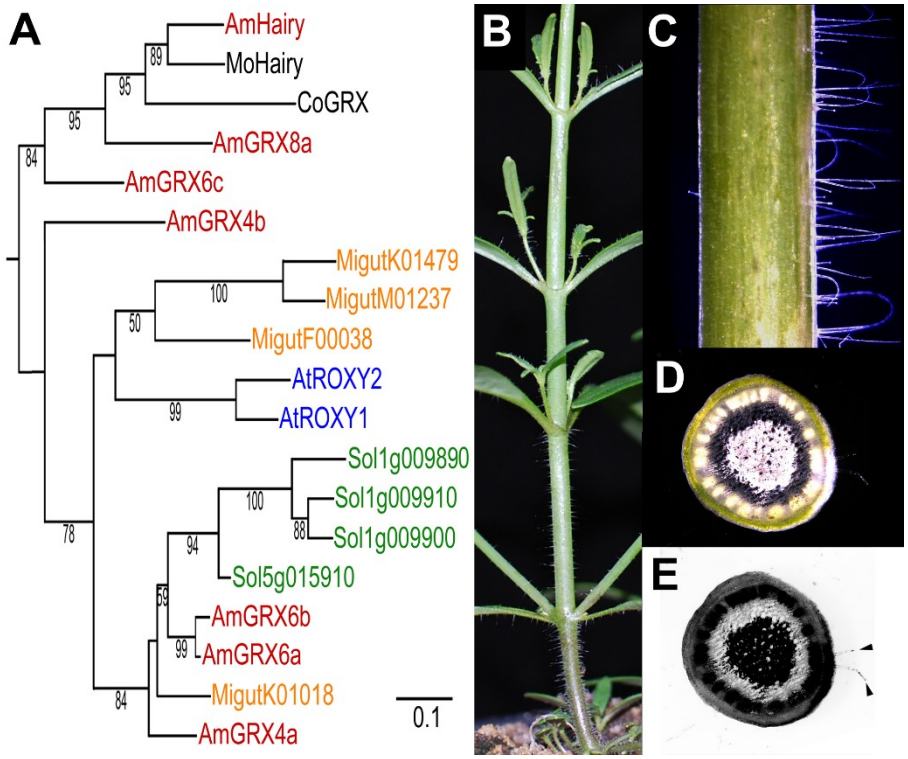


Figure 5

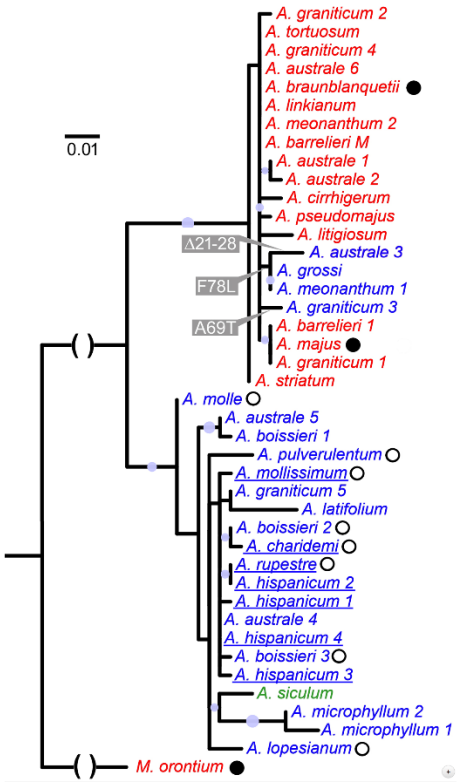


Figure 6

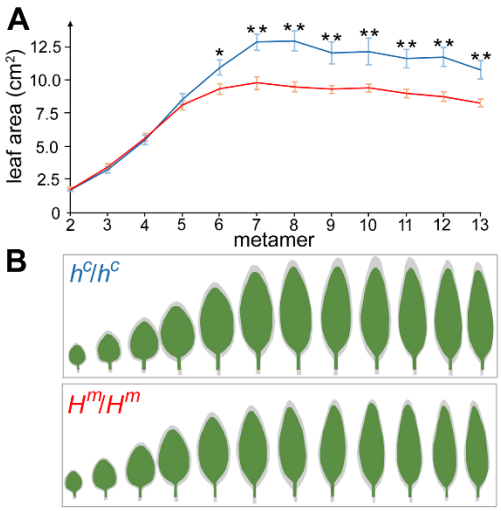

A
leaf area (cm²)
12.5
10.0
7.5
5.0
2.5
0
2
3
4
5
6
7
8
9
10
11
12
13
metamer
*
**
**
**
**
**
**
**
B
h^c/h^c
H^m/H^m

Figure 7

SUPPLEMENTAL INFORMATION

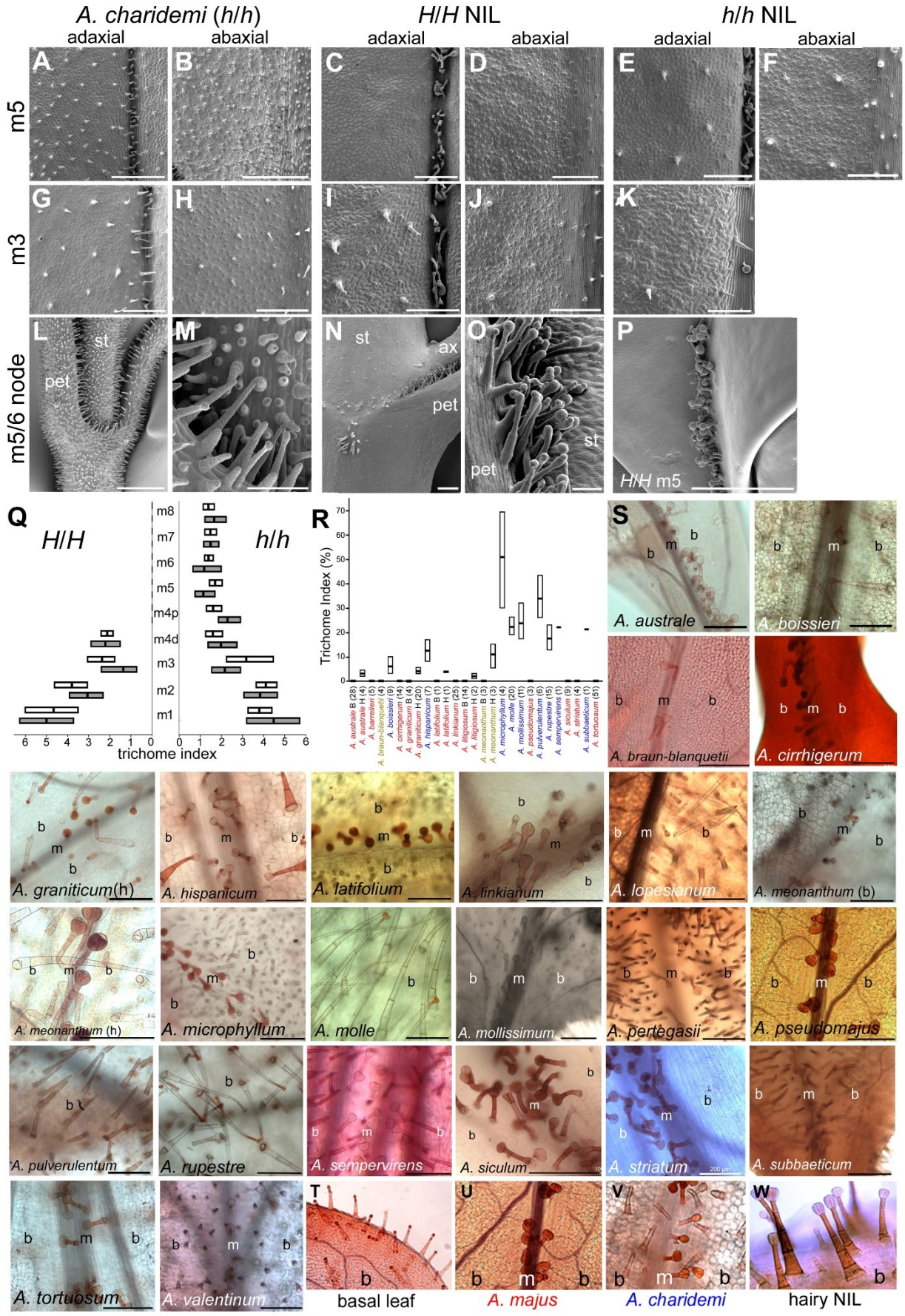


Figure S1 Details of trichome phenotypes, related to Figure 1

(A-K and P) Scanning electron micrographs of the adaxial or abaxial surface of leaves from *A. charidemi*, the bald (H^m/H^m) NIL or the hairy (h^c/h^c) NIL. Midribs are to the right of images A-K). In the H^m/H^m NIL, the abaxial epidermis of leaves above m4 (here shown for m5) lacks trichomes (D), while in the adaxial epidermis (C and P), trichomes are confined to the midrib. In basal leaves of the h^c/h^c NIL (not shown), the distribution of trichomes is indistinguishable from that in basal leaves of the H^m/H^m NIL shown for m3 in (J). (L,M) The stem (st) and abaxial petiole (pet) of *A. charidemi* above m4 are covered in glandless trichomes, while glandular trichomes (arrowheads in M) are found on the adaxial midrib and the junction of the adaxial petiole with the stem. The same distribution of glandular trichomes is seen in the NILs and shown here for the m5-6 node of the bald (H^m/H^m) NIL in N-O) (ax, axillary shoot). All scale bars are 0.5 mm, except in M) and O) where they represent 0.1 mm. Q) The percentage of epidermal pavement cells that carry trichomes (trichome index) in the leaf blades of representative bald (*H*/*H*) or hairy (*h*/*h*) NILs (adaxial grey boxes, abaxial unfilled). In *H*/*H* plants, the transition from hairy to bald occurred part-way along m4 leaves, therefore separate values are given for the distal, earlier maturing, part of the m4 leaf (m4d) and for the proximal, later maturing part (m4p). Each box represents the mean for ten areas sampled from two plants, with 95% confidence intervals estimated by bootstrapping. No trichomes were observed on leaf blades of the *H*/*H* NIL above m4, only on the adaxial midrib, as shown in (C, P). All plants were grown in the same conditions (see STAR Methods). R) Adaxial trichome indices in m5 leaves of different species (mean ± SE for the sample size shown in brackets). Hairy (H) and bald (B) phenotypes in polymorphic species are shown separately. S) Micrographs adaxial surfaces of m5 leaves from different species, usually including both leaf blade (b) and midrib (m). All scale bars are 200 µm. T) Both *A. majus* and *A. charidemi* produce glandular secretory trichomes from the blade (b) and midrib (m) of basal leaves (below m5). From m5 onwards *A. majus* leaves (U) make trichomes only from the adaxial midrib (m), and leaf blades are bald. At the same nodes, *A. charidemi* (V) has the same type of glandular trichomes on the midrib as *A. majus*, however it produces short, non-secretory trichomes from the leaf blade. The hairy NIL, homozygous for the *A. charidemi* h^c allele, produces trichomes on the blades of leaves from m5 onwards (W), as does its *A. charidemi* parent. However, these trichomes are glandular, unlike those of *A. charidemi*, indicating that the *Hairy* gene underlies variation in trichome distribution, but not trichome morphology. Although unicellular, non-secretory trichomes have been reported in *A. charidemi* [S1], they were possibly immature because only multicellular trichomes were observed in the mature tissues studied here.

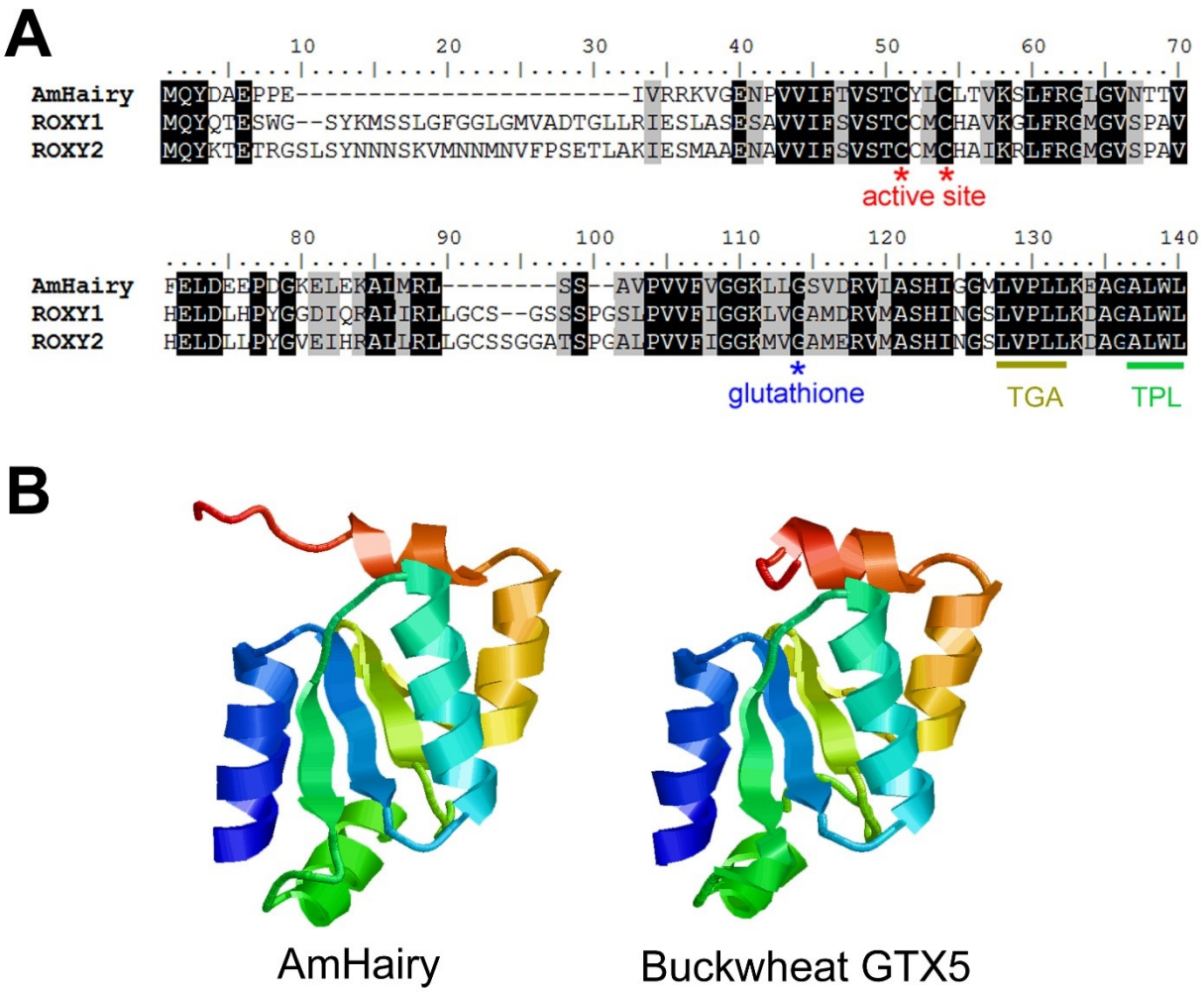


Figure S2 *Hairy* encodes a CC-family glutaredoxin, related to Figure 2

A) Clustal alignment of the predicted full-length Hairy protein from *A. majus* with the most similar proteins from *Arabidopsis thaliana*–ROXY1 (ROX1, the product of At3g02000) and ROX2 (At5g14070). Identical residues are boxed in black, conservative substitutions in grey. The two catalytically active cysteines, conserved among almost all other CC-family glutaredoxins, and the glycine residue required for glutathione cofactor binding, are highlighted. Although Hairy is nested within the CC-family (Figure S6A), it lacks the second of the two adjacent active-centre cysteins (CC in red type), which are ancestral in the family, retained by most other members [S2], and give the family its name. The hydrophobic L**LL motif, required by ROX1 and ROX2 for interaction with TGA transcription factors [S3], and the ALWL motif required for interaction with transcriptional co-repressors of the TPL family [S4] are also shown. All the motifs are also present in the *H* paralogue *GRX6c* (not shown). B) Homology-informed secondary structure prediction for *A. majus* Hairy, made with Phyre2 [S5], compared to the experimentally determined structure of a GRX from buckwheat, expressed in *E. coli* [S6].

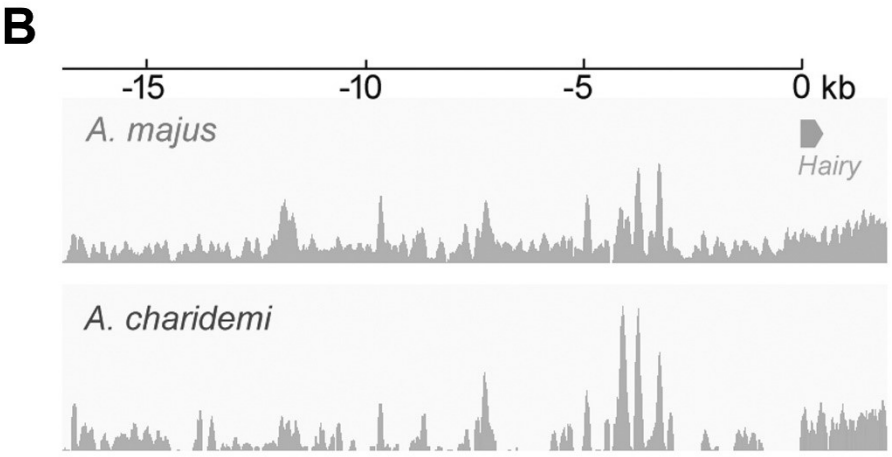


Figure S3 Comparison of *A. majus* and *A. charidemi GRX8b* (*H*) haplotypes, related to Figures 2 and 4

A) Alignment of DNA sequences (*A. majus* and *A. charidemi*) and encoded proteins (A. majus and A. charidemi). Identical nucleotides or amino acids are boxed grey. B) Coverage of unfiltered sequencing reads from *A. majus* or *A. charidemi* mapping in the region 5' to *H* in the *A. majus* reference genome, which contains no predicted genes. Repetitive regions (e.g., transposons) give rise to peaks in read coverage. Comparison of coverage between *A. majus* and *A. charidemi* suggests multiple insertion/deletion polymorphisms in the *Hairy* promoter region.

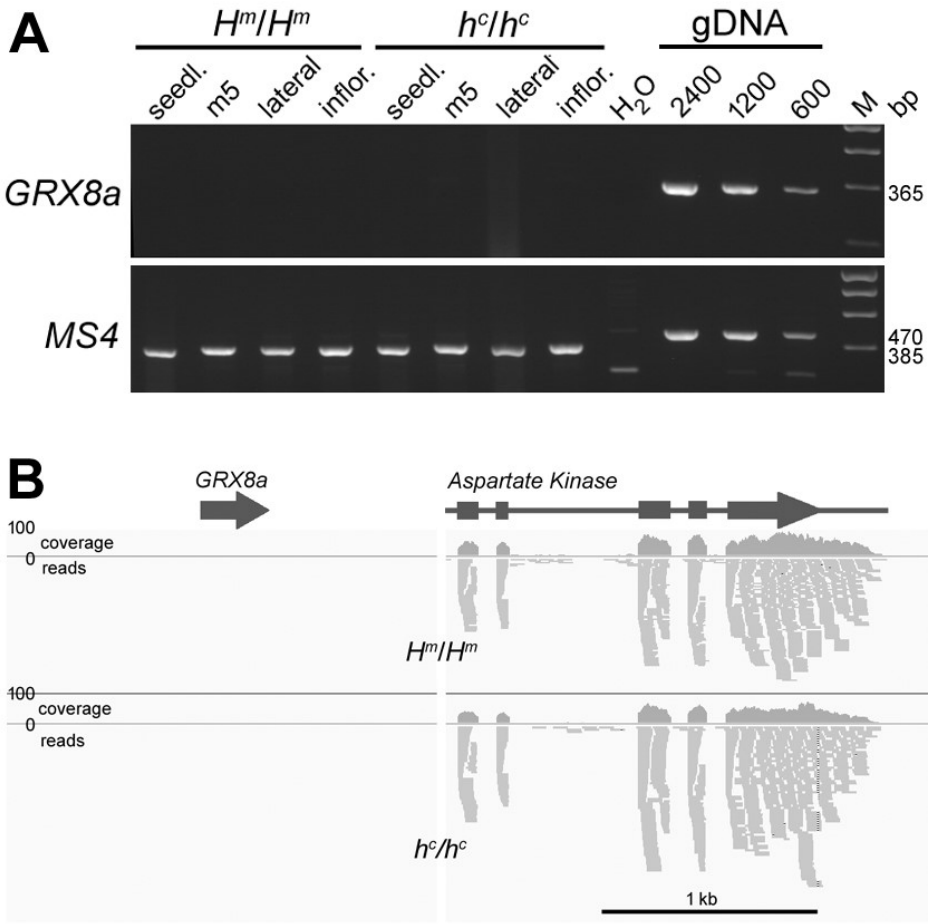


Figure S4 Testing expression of the *H* paralogue *GRX8a*, related to Figure 2

A) cDNA was made from different aerial tissues of bald NILs (H^m/H^m), homozygous for *A. majus* alleles of both *H* and its closely-linked paralogue *GRX8a*, or hairy NILs (h^c/h^c), homozygous for *A. charidemi* alleles at both loci. Tissues were aerial parts of whole seedlings (seedl.), vegetative shoot tips with 5 mm leaves at metamer 5 (m5), tips of vegetative axillary shoots from above m5 (lateral) and inflorescence apices (infl.). Expression of *GRX8a* could not be detected by RT-PCR using primers within its single predicted exon, in either the presence or absence of H activity, though as few as 600 copies of *GRX8a* could be detected in a genomic DNA (gDNA) template under the same conditions. (This copy-number estimate assumes that all gDNA is nuclear and is therefore conservative.) cDNA from the housekeeping gene *Methionine synthase 4* (*MS4*) was detectable in all samples. M is a size marker. *MS4* primers span an 85 bp intron, therefore amplification from gDNA gives a larger (470 bp) product. A small product (<385 bp) amplifies from bacterial DNA contamination of *Taq* polymerase in the absence of added template (H_2O). B) RNAseq reads from vegetative apices of bald (H^m/H^m) or hairy (h^c/h^c) NILs, mapping either to the *GRX8a* region or to a linked aspartate kinase-encoding gene. The broad arrow shows the *GRX8a* ORF, predicted by homology.

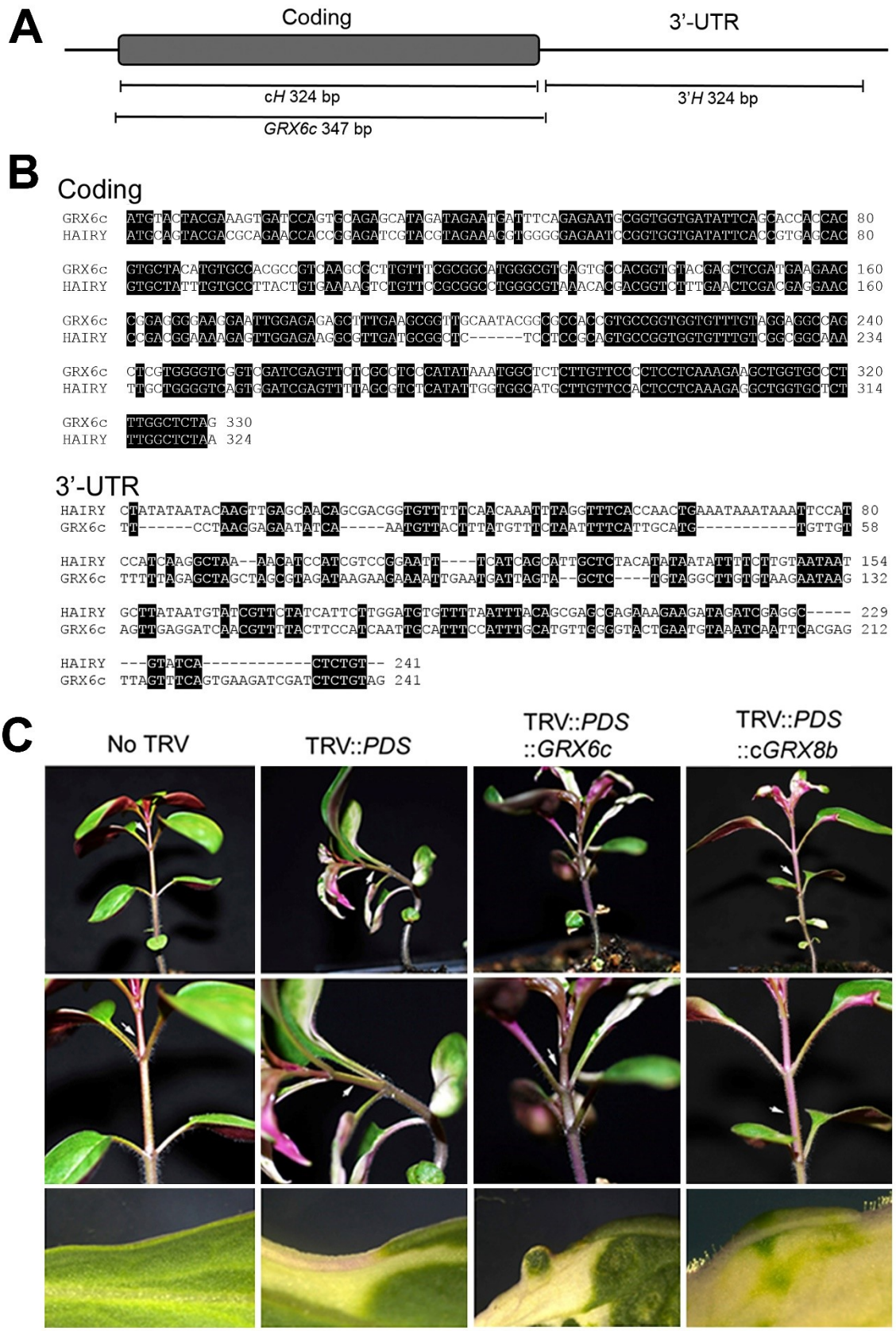


Figure S5 ViGS of *H* and its paralogue, *GRX6c*, related to Figure 3 and STAR Methods

A) ViGS of *H* was achieved either with the whole coding region (c*H*) or most of the 3'-UTR (3'*H*), attempted ViGS of the *H* paralogue *GRX6c* used a sequence extending from 2 bp upstream of the initiation codon to 38 bp downstream of the Stop codon. Both genes lack introns. B) Sequence alignments of the coding or 3'-UTRs of *H* and *GRX6c*. The longest block of identical sequence is 17 bp. C) ViGS using a parts of the *GRX6c* and *PDS* sequences reduced *GRX6c* RNA abundance to below 20% of normal levels (Figure 3G), but did not affect trichome development in bleached tissue showing ViGS (third column). In contrast, ViGS with part of the coding sequence of *H* (bottom row) allowed formation of ectopic trichomes. The bottom row shows parts of representative metamer 5 leaves. Arrows show the position of the last vegetative stem trichome, which is also unaffected by reduced *GRX6c* expression.

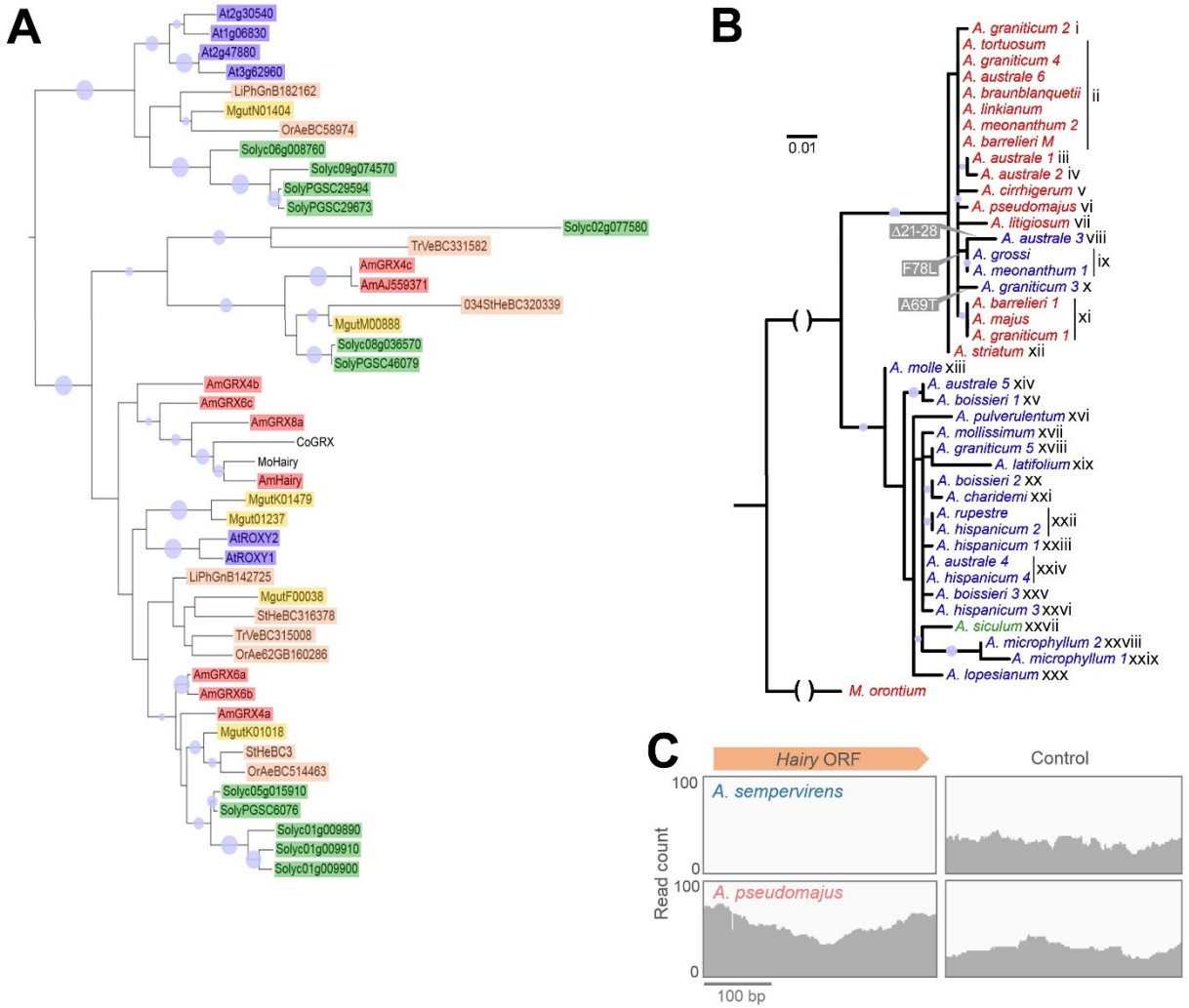


Figure S6 Relationships of Hairy-like GRX proteins and *H* haplotypes, related to Figures 5 and 6

A) The most probably ML tree of H-like proteins from selected eudicot species was obtained using RaxML 8.0 with the same parameters as those used for the subclade containing H in Figure 5A. Sequences from the Orobanchaceae species *Orobanche aegyptiaca* (OrAe), *Striga hermontheca* (StHe), *Triphysaria versicolor* (TrVe) and *Lindenbergia philippensis* (LiPh), which are coloured carmine, were translated from EST databases of the Parasitic Plant Genome Project (http://ppgp.huck.psu.edu/). They were not used for the H subtree in Figure 5A because of the possibility of other paralogues remaining unsampled as ESTs. The remaining sequences from *A. majus* (red), Mimulus (yellow), tomato (green) and Arabidopsis (blue) are from annotated reference genomes (Table S5). Nodes recovered in ≥50% of 520 bootstrap replicates have circles of a size proportional to the level of support. B) Distribution of *H* haplotypes among species. *Hairy* coding sequences are named according to the species from which they originated, with Arabic numerals used to distinguish different sequences from the same species (e.g., *A. microphyllum* 1 and *A. microphyllum* 2). Because more than one species could share the same sequence, each unique sequence (haplotype) was given a number in Roman numerals (*e.g.*, haplotype *xi* was found in both *A. grossi* and *A. meonanthum*). The distribution of sampled haplotypes among species and populations is summarised in Table S4. C) Evidence for deletion of *H* from members of the

basal *Kickxiella* lineage. Coverage of genome resequencing reads at *Hairy* (left) or an unlinked single-copy region (right). The reads are from *A. sempervirens*, a member of the basal *Kickxiella* lineage from which *H* could not be amplified, or *A. pseudomajus* (genotype *H*/*H* from subsection *Antirrhinum*). Reads were mapped with the same parameters in both cases.

Trait	**Hairy** (H^m/H^m)	**Bald** (h^c/h^c)	***p*-value†**
First spiral node	11.8 (±0.4), *n*=13	12.4 (±0.6), *n*=16	0.37
First flowering metamer*	17.8 (±1.4), *n*=13	19.8 (±1.0), *n*=16	0.25
Frequency of adaxial midrib trichomes (mm^{-1})	9.31 (±0.40), n=20	8.71 (±0.51), n=20	0.36

Table S1 Traits related to developmental phase and the frequency of midrib trichomes in hairy and bald NILs. Related to STAR Methods.

Values are means for hairy plants or bald plants in a near-isogenic background ± their standard errors. * first flowering metamer is equivalent to the number of metamers that produced leaves before flowering and is therefore a proxy for flowering time; † probabilities from Students *t*-tests of means being the same.

Number*	**Most similar in *A. thaliana*†**	**Annotation *A. thaliana* Araport11**	**Expression in *Antirrhinum* ‡**
1	At5g14020	Unknown function. Contains endosomal targeting BRO1-like domain.	not detected
2	At5g41080	Member of the glycerophosphodiester phosphodiesterase (GDPD) family.	15
3	At5g50810	Inner mitochondrial membrane translocase subunit TIM8.	60
4	At5g18610	Receptor-like cytoplasmic kinase immediately downstream of the CIRK1 chitin receptor in chitin-induced immunity	1.2
5	At5g14070	ROXY2. CC-type glutaredoxin that controls anther development with ROXY1.	not detected
6	At5g14070	ROXY2, as above.	8.2

Table S2 Genes in the genome interval containing *Hairy*. Related to Figure 2B.

* *Antirrhinum* genes are numbered as in Figure 2B; † The gene encoding the product with the lowest E-value in BlastX searches of Araport11; ‡ Transcripts Per Million (TPM) values for shoot apices of the *Antirrhinum* H^m/H^m NIL.

Male parent	Subsection*	Parent phenotype	N bald F1	N hairy F1	*p*(1:1)†
A. majus	*A*	Bald	68	0	-
A. linkianum	*A*	Bald	30	0	-
A. pseudomajus	*A*	Bald	31	0	-
A. striatum	*A*	Bald	37	0	-
A. tortuosum	*A*	Bald	63	0	-
A. siculum	*A*	Bald	0	1‡	-
A. hispanicum	*K*	Hairy	17	14	0.59
A. boissieri	*K*	Hairy	16	21	0.41
A. graniticum	*A*	Hairy	20	22	0.76
A. lopesianum	*K*	Hairy	6	6	1.00
A. molle	*K*	Hairy	19	15	0.49
A. rupestre	*K*	Hairy	9	11	0.65
A. pulverulentum	*K*	Hairy	27	22	0.48
A. subbaeticum	*K*	Hairy	7	6	0.78
A. valentinum	*K*	Hairy	19	17	0.74

Table S3 Allelism tests. Related to STAR Methods.

Each species was crossed as pollen parent to the NIL that was heterozygous H^m/h^c and the numbers of hairy and bald phenotypes in the progeny recorded. * *A*, subsection *Antirrhinum*, *K*, subsection *Kickxiella*; † *p*-values from χ^2-tests fitting the observed numbers of F1 phenotypes to the 1:1 ratio expected if the hairy parent was homozygous *h*/*h*; ‡ few viable seeds were produced by crossing *A. siculum* to the NIL.

Subsection	Species	Code	Location	Haplotype
Antirrhinum	*A. australe* Rothm.	L084*	El Boyar, Cádiz, Spain	xiv
		L091*	Gaucín, Málaga, Spain	ii, xiv
		L095*	El Burgo, Málaga, Spain	ii, iii, iv, viii, xxiv
	A. barrelieri Boreau	L150	Cádiar, Granada, Spain	xi
	"A. barrelieri"	L167	Taineste, Morocco	ii
	A. graniticum Rothm.	L040*	Bragança, Bragança, Portugal	i, ii, x, xi, xvii
		L069	Ledanca, Guadalajara, Spain	xi
		L116*	Celorico da Beira, Guarda, Portugal	xxvii
	A. latifolium Miller	AC1066	St Martin d'Entraunes, Alpes-Maritimes, Fra.	xix
	A. cirrhigerum Filcaho	L114	Praia de Mira, Coimbra, Portugal	v
	A. linkianum Boiss & Reuter	L108	Almada, Setúbal. Portugal	ii
	A. litigiosum Pau	L003	Cheste, Valencia, Spain	xi
	A. majus L.		JI.7 from John Innes Centre	xi
	A. pseudomajus Rouy	L053	Minerve, Hérault, France	vi
	A. striatum Rothm.	AC1125	Alet-les-Bains, Aude, France	iii
	A .tortuosum Rouy	L092	Casares, Málaga, Spain	ii
	A. siculum Millar	AC1177	Taormina, Sicily, Italy	xxvii
Kickxiella	*A. boissieri* Rothm.	L018	Guadahortuna, Granada, Spain	xv, xxv
		L104	Ermita Virgen de la Sierra, Cordoba, Spain	xx
	A. charidemi Lange	E023	Cabo de Gata, Almería, Spain	xxi
	A. grosii Font Quer	L175	Sierra de Gredos, Toledo, Spain	ix
	A. hispanicum Chav.	L030	Mecina Bombarón, Granada, Spain	xxii, xxiii
		L036	Balcón de Canales, Granada, Spain	xxiv, xxvi
	A. lopesianum Rothm.	L038	Vimioso, Bragança, Portugal	xxx
	A. microphyllum Rothm.	L072	Sacedón, Guadalajara, Spain	xxix
		L073	Pantano de Buendía, Guadalajara, Spain	Δ†
		L074	Buendía, Guadalajara, Spain	xxviii
	A. molle Lange	E051	Gerri de la Sal, Lleida, Spain	xii
	A. mollisimum Rothm.	L021	Enix, Almería, Spain	xvii
	A. pertegasii Rothm.	E065	Castellón de la Plana, Castellón, Spain	Δ
	A. pulverulentum Lazaro	L068	Pelegrina, Guadalajara, Spain	Δ
		L070	Alcorlo, Guadalajara, Spain	xvi
		L077	Poveda de la Sierra, Guadalajara, Spain	Δ
	A. rupestre Rothm.	L139	Capileira, Granada, Spain	xxii
	A. sempervirens Lapeyre	L050	Col d'Aubisque, Hautes-Pyrenees, France	Δ
		L052	Luz-Saint-Sauveur, Hautes-Pyrenees, France	Δ
	A. subbaeticum Guemes	E72	Albacete, Albacete, Spain	Δ
	A. valentinum Font Quer	AC1173	La Drova, Valencia, Spain	Δ
Streptosepalum	*A. meonanthum* Hoffmans & Link	L118	Manteigas, Guarda, Portugal	ii, ix
	A. braun-blanquetii Rothm.	E20	St Pietro de Villanueva, Asturias, Spain	ii
Other species	*Chaenorhinum origanifolium* Fourr.		Gijón, Asturias, Spain	*CoGRX*
	Misopates orontium Raf.		Unknown	*MoHairy*

Table S4 Accessions of *Antirrhinum* and related species. Related to STAR Methods.

* polymorphic population with both bald and hairy individuals, † Representatives of endemic *Kickiella* species in which the *H* locus could not be detected and is presumed to have been deleted are shown with Δ. For relationships of haplotypes (Roman numerals) please see Figure S6B.

***Arabidopsis thaliana* TAIR10**	
AtROXY1	At3g02000
AtROXY2	At5g14070
***Mimulus guttatus* v2.0**	
MigF00038	Migut.F00038
MigK01018	Migut.K01018
MigK01479	Migut.K01479
MigM01237	Migut.M01237
***Solanum esculentum* iTAG2.4**	
S01g009890	Soly01g009890
S01g009900	Soly01g009900
S01g009910	Soly01g009910
S05g015910	Soly05g015910

Table S5 Origins of GRX peptide sequences. Related to STAR Methods.

Use	Gene	Primer Name	Sequence (5'-3')
Genotyping	*Antirrhinum FLORICAULA*	FLO-F	GGAAGTGAGGCGGAGGCA
		FLO-R	ACCCGCCCCCATCATTC
	Antirrhinum S-linked F-box	SLF-F	GTGCTTTCCTTCCACGATGT
		SLF-R	CCTGGTTCAAACTGATCAAGC
	Antirrhinum CYCLOIDEA	CYC-F	TCCTCCCTTCACTCTCGCGC
		CYC-R	TGGCGCATAGCTGGTTCGAC
RT-PCR	*A. majus H*	AmHQ-F	TCTTGTCTTTTCCACCTGTCA
		AmHQ-R	TGAATATCACCACCGGATTCTC
	A. majus GRX6c	AmGRX6cQ-F	TGGCTCTAGTTCCTAAGGAGAA
		AmGRX6cQ-R	CACAAGCCTACAGAGCTACTAATC
	A. majus PDS	AmPDSQ-F	TCTTTGTAATGGACGGCAAG
		AmPDSQ-R	ACTTGCCAAACTCTTCCCTG
	A. majus Ubi5	AmUbiQ-F	CCGAACCATCAGACAAACAAAC
		AmUbiQ-R	TACCCTGGCCGACTACAATA
	A. majus MS4	AmMS4-F	GTTTGATGAGCCCACCCTTG
		AmMS4-R	TGTGGAGAAGTGAGCAGGAG
	A. majus GRX8a	AmGRX8a-F	CACCCATCAATTAACTTCACCAAAATG
		AmGRX8a-R	ATCACAAGCATTAGAGCCAAAG
Phylogenetic analysis	*Antirrhinum H*	AmH-R	GTAGTCCTATACAAATTAATACGTA
		AmH-F	ACAGAGTATACGCCTCGAT
		AmH-F3	GTTTCCCTGGAATCAACCAC
		AmH-R3	AACACCGTCGCTGTTGCTC
		AmH-R2	ACAGTCCTATACAAATTAATATG
	Misopates & *Chaenorrhinum GRX*	GRX-RACE-R3	ACCAAGTGGCGTACGAAATTA
		GRX-CR-R	GAATCCGGTGGTGATATTCA

Table S6 Oligonucleotides used in this study. Related to STAR Methods.

Supplemental Reference List

[S1] Doaigey, A.R. and Harkiss, K.J. (1991). Application of epidermal characters to the taxonomy of European species of Antirrhinum (Schrophulariaceae). Nord. J. Bot. *11*, 513-524.

[S2] Gutsche, N., Holtmannspötter, M., Maß, L., O'Donoghue, M., Busch, A., Lauri, A., Schubert, A. and Zachgo, S. (2017). Conserved redox-dependent DNA binding of ROXY glutaredoxins with TGA transcription factors. Plant Direct *1*, 1-17.

[S3] Li, S., Gutsche, N. and Zachgo, S. (2011). The ROXY1 C-terminal L**LL motif is essential for the interaction with TGA transcription factors. Plant Physiol. *157*, 2056-2068.

[S4] Uhrig, J.F., Huang, L.J., Barghahn, S., Willmer, M., Thurow, C. and Gatz, C. (2017). CC-type glutaredoxins recruit the transcriptional co-repressor TOPLESS to TGA-dependent target promoters in *Arabidopsis thaliana*. Biochim. Biophys. Acta *1860*, 218-226.

[S5] Kelly, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2012). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protoc. *10*, 845-858.

[S6] Zhang, X., Wang, W., Li, C., Zhao, Y., Yuan, H., Tan, X., Wu, L., Wang, Z. and Wang, H. (2017). Structural insights into the binding of buckwheat glutaredoxin with GCH and regulation of its catalytic activity. J. Inorg. Biochem. *173*, 21-27.